Title: A NOVEL BOTULINUM NEUROTOXIN AND ITS DERIVATIVES

Abstract: Provided herein are Clostridial Botulinum neurotoxin (BoNT) polypeptides of a novel serotype (BoNT/X) and methods of making and using the BoNT polypeptides, e.g., in therapeutic applications.
Published:
— without international search report and to be republished upon receipt of that report (Rule 48.2(g))
— with sequence listing part of description (Rule 5.2(a))
A NOVEL BOTULINUM NEUROTOXIN AND ITS DERIVATIVES

RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. provisional application number 62/360,239, filed July 8, 2016, which is incorporated by reference herein in its entirety.

GOVERNMENT SUPPORT

This invention was made with government support under R01NS080833 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND

*Clostridial Botulinum* neurotoxins (BoNTs) are among the most dangerous potential bioterrorism agents and are also used clinically to treat a growing list of medical conditions. There are seven serotypes of BoNTs (BoNT/A-G) known to date. In recent years, BoNTs have been widely used to treat a growing list of medical conditions: local injections of minute amount of toxins can attenuate neuronal activity in targeted regions, which can be beneficial in many medical conditions as well as for cosmetic purposes. As the application of BoNTs grows, limitations and adverse effects have been reported. The major limitation is the generation of neutralizing antibodies in patients, which renders future treatment ineffective. Termination of BoNT usage often leaves patients with no other effective ways to treat/relieve their disorders. Adverse effects associated with BoNT use range from transient non-serious events such as ptosis and diplopia to life-threatening events even death. The limitations and adverse effects of BoNTs are largely correlated with dose. There are considerable interests in developing novel BoNT types as therapeutic toxins. No new BoNT types have been recognized for the past 45 years.

SUMMARY

The present disclosure is based, at least in part, on the identification of a novel BoNT serotype, BoNT/X, from searching genomic database of *Clostridium Botulinum* strains. BoNT/X the lowest sequence identity with other BoNTs and it is not recognized by antisera raised against known BoNT types. BoNT/X cleaves SNARE proteins, like other BoNTs.
However, BoNT/X also cleave several SNARE proteins that other BoNTs cannot cleave, e.g., VAMP4, VAMP5, and Ykt6. Compositions and methods for treating diseases using BoNT/X are provided. Also provided herein are methods of making BoNT/X.

Accordingly, some aspects of the present disclosure provide isolated Clostridial Botulinum neurotoxin (BoNT) polypeptides comprising the amino acid sequence of SEQ ID NO: 1.

Some aspects of the present disclosure provide isolated BoNT polypeptides comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 1. In some embodiments, the isolated BoNT polypeptide consists of the amino acid sequence of SEQ ID NO: 1.

Some aspects of the present disclosure provide isolated BoNT polypeptides comprising the amino acid sequence of SEQ ID NO: 2. Some aspects of the present disclosure provide isolated BoNT polypeptides an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 2. In some embodiments, the isolated BoNT polypeptide consists of the amino acid sequence of SEQ ID NO: 2.

Some aspects of the present disclosure provide isolated BoNT polypeptides comprising the amino acid sequence of SEQ ID NO: 3. Some aspects of the present disclosure provide isolated BoNT polypeptides an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 3. In some embodiments, the isolated BoNT polypeptide consists of the amino acid sequence of SEQ ID NO: 3.

In some embodiments, the modified BoNT polypeptide comprises the amino acid sequence of any one of SEQ ID NO: 4-17. In some embodiments, the modified BoNT polypeptide comprises an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99.5% identity to any of SEQ ID NOs: 4-17, wherein the polypeptide does not have the amino acid sequence of SEQ ID NO: 1. In some embodiments, the modified BoNT polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 4-17.

Some aspects of the present disclosure provide modified BoNT polypeptides comprising a single substitution mutation in a position corresponding to C461 or C467 of SEQ ID NO: 2.


In some embodiments, the modified BoNT polypeptide comprises the amino acid sequence of any one of SEQ ID NOs: 18-21. In some embodiments, the modified BoNT polypeptide comprises an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any of SEQ ID NOs: 18-21, wherein the polypeptide does not have the amino acid sequence of SEQ ID NO: 2. In some embodiments, the modified BoNT polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 18-21.

Some aspects of the present disclosure provide chimeric BoNT polypeptides comprising the amino acid sequence of any one of SEQ ID NOs: 22-24. In some embodiments, the chimeric BoNT polypeptide comprises an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any one of SEQ ID NOs: 22-24, wherein the polypeptide does not have the amino acid sequence of SEQ ID NO: 1 or SEQ ID NO: 2. In some embodiments, the chimeric BoNT polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 22-24.
In some embodiments, the chimeric BoNT polypeptide further comprises a single substitution mutation in a position corresponding to C461 or C467 of in SEQ ID NO: 2.

In some embodiments, the chimeric BoNT polypeptide comprises the amino acid sequence of any one of SEQ ID NOs: 25-30. In some embodiments, the chimeric BoNT polypeptide comprises an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any one of SEQ ID NOs: 25-30. In some embodiments, the chimeric BoNT polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 25-30.

In some embodiments, the BoNT polypeptide enters a cell. In some embodiments, the BoNT polypeptide cleaves a SNARE protein in the cell. In some embodiments, the SNARE protein is selected from the group consisting of: SNAP-25, VAMP1, VAMP2, VAMP3, VAMP4, VAMP5, Ykt6, and syntaxin 1.

In some embodiments, the SNARE protein is VAMP1. In some embodiments, the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 39.

In some embodiments, the SNARE protein is VAMP2. In some embodiments, the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 40.

In some embodiments, the SNARE protein is VAMP3. In some embodiments, the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 41.

In some embodiments, the SNARE protein is VAMP4. In some embodiments, the BoNT cleaves between amino acid residues corresponding to K87 and S88 of SEQ ID NO: 42.

In some embodiments, the SNARE protein is VAMP5. In some embodiments, the BoNT cleaves between amino acid residues corresponding to R40 and S41 of SEQ ID NO: 43.

In some embodiments, the SNARE protein is Ykt6. In some embodiments, the BoNT cleaves between amino acid residues corresponding to K173 and S174 of SEQ ID NO: 44.

In some embodiments, the BoNT polypeptide has increased stability compared to its corresponding wild type BoNT polypeptide.

In some embodiments, the cell is a secretory cell. In some embodiments, the cell is a neuronal cell. In some embodiments, the cell is an immune cell. In some embodiments, the BoNT polypeptide suppresses neuronal activity. In some embodiments, the BoNT polypeptide induces flaccid paralysis. In some embodiments, the cell is a cultured cell. In some embodiments, the cell is from a mammal. In some
embodiments, the mammal is a human. In some embodiments, mammal is a rodent. In some
embodiments, the rodent is a mice. In some embodiments, the rodent is a rat.

In some embodiments, the BoNT polypeptide does not cross react with an antibody
against BoNT serotype A, B, C, D, E, F, or G.

Other aspects of the present disclosure provide nucleic acid molecules comprising a
polynucleotide encoding a polypeptide comprising an amino acid sequence that has at least
85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least
92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least
99%, or at least 99.5%, or 100% identity to the BoNT polypeptide described herein. Nucleic
acid vectors comprising such nucleic acid molecules are provided. Cells comprising the nucleic
acid molecules or the nucleic acid vectors described herein are provided. In some
embodiments, such cells express the BoNT polypeptide described herein.

Methods of producing the BoNT polypeptide of the present disclosure are provided.
Such methods comprise the steps of culturing the cell expressing the BoNT polypeptides under
conditions wherein said BoNT polypeptide is produced. In some embodiments, the methods
further comprise recovering the BoNT polypeptide from the culture.

Other aspects of the present disclosure provide modified BoNT polypeptides
comprising: (a) a protease domain; (b) a modified linker region; and (c) a translocation
domain; wherein (a), (b), and (c) are from BoNT serotype X, and wherein the modified linker
region comprises one single substitution mutation in a position corresponding to C461 or C467
of SEQ ID NO: 1.

In some embodiments, the modified BoNT polypeptide further comprises: (d) a
receptor binding domain.

In some embodiments, modified linker region comprises a substitution mutation
corresponding to C461S or C461A in SEQ ID NO: 1. In some embodiments, the modified
linker region comprises a substitution mutation corresponding to C467S or C467A in SEQ ID
NO: 1.

In some embodiments, the receptor binding domain is from BoNT/X. In some
embodiments, the receptor binding domain is modified. In some embodiments, the receptor
binding domain comprises a substitution mutation corresponding to C1240S or C1240A in
SEQ ID NO: 1.

In some embodiments, the receptor binding domain is from a serotype selected from the
group consisting of A, B, C, D, E, F, and G.
In some embodiments, the modified BoNT polypeptide enters a cell. In some embodiments, the modified BoNT polypeptide cleaves SNARE proteins in the cell. In some embodiments, the SNARE protein is selected from the group consisting of: SNAP-25, VAMP1, VAMP2, VAMP3, VAMP4, VAMP5, Ykt6, and syntaxin 1.

In some embodiments, the SNARE protein is VAMP1. In some embodiments, the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 39.

In some embodiments, the SNARE protein is VAMP2. In some embodiments, the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 40.

In some embodiments, the SNARE protein is VAMP3. In some embodiments, the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 41.

In some embodiments, the SNARE protein is VAMP4. In some embodiments, the BoNT cleaves between amino acid residues corresponding to K87 and S88 of SEQ ID NO: 42.

In some embodiments, the SNARE protein is VAMP5. In some embodiments, the BoNT cleaves between amino acid residues corresponding to R40 and S41 of SEQ ID NO: 43.

In some embodiments, the SNARE protein is Ykt6. In some embodiments, the BoNT cleaves between amino acid residues corresponding to K173 and S174 of SEQ ID NO: 44.

In some embodiments, the BoNT polypeptide has increased stability compared to its corresponding wild type BoNT polypeptide.

In some embodiments, the cell is a secretory cell. In some embodiments, the cell is a neuronal cell. In some embodiments, the cell is an immune cell. In some embodiments, the BoNT polypeptide suppresses neuronal activity. In some embodiments, the BoNT polypeptide induces flaccid paralysis. In some embodiments, the cell is a cultured cell. In some embodiments, the cell is in vivo. In some embodiments, the cell is from a mammal. In some embodiments, the mammal is a human. In some embodiments, mammal is a rodent. In some embodiments, the rodent is a mice. In some embodiments, the rodent is a rat.

In some embodiments, the BoNT polypeptide does not cross react with an antibody against BoNT serotype A, B, C, D, E, F, or G.

In some embodiments, the modified linker region comprises an artificial linker. In some embodiments, the artificial linker contains a cleavage site of a protease. In some embodiments, the protease is selected from the group consisting of Thrombin, TEV, PreScission (3C protease), Factor Xa, MMP-12, MMP-13, MMP-17, MMP-20, Granzyme-B, and Enterokinase. In some embodiments, the linker comprises the amino acid sequence of any of SEQ ID NOs: 50-60).
Other aspects of the present disclosure provide nucleic acid molecules comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%, or 100% identity to the BoNT polypeptide described herein. Nucleic acid vectors comprising such nucleic acid molecules are provided. Cells comprising the nucleic acid molecules or the nucleic acid vectors described herein are provided. In some embodiments, such cells express the BoNT polypeptide described herein.

Methods of producing the BoNT polypeptide of the present disclosure are provided. Such methods comprise the steps of culturing the cell expressing the BoNT polypeptides under conditions wherein said BoNT polypeptide is produced. In some embodiments, the methods further comprise recovering the BoNT polypeptide from the culture.

Other aspects of the present disclosure provide modified BoNT polypeptides comprising one or more substitution mutation(s) in positions corresponding to R360, Y363, H227, E228, or H231 in SEQ ID NO: 1. In some embodiments, the one or more substitution mutation corresponds to R360A/Y363F, H227Y, E228Q, or H231Y in SEQ ID NO: 1. In some embodiments, the modified BoNT polypeptide comprises the amino acid sequence of any one of SEQ ID NOs: 31-38. In some embodiments, the modified BoNT polypeptide comprises an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any of SEQ ID NOs: 31-38, wherein the polypeptide does not have the amino acid sequence of SEQ ID NO: 1 or SEQ ID NO: 2. In some embodiments, the modified BoNT polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 31-38.

Other aspects of the present disclosure provide modified BoNT/X polypeptide comprising: a) an inactive protease domain; b) a linker region; and c) a translocation domain. In some embodiments, the modified BoNT/X further comprises a receptor binding domain.

In some embodiments, the inactive protease domain comprises one or more substitution mutations in positions corresponding to R360, Y363, H227, E228, or H231 of SEQ ID NO: 1. In some embodiments, the one or more substitution mutations correspond to R360A/Y363F, H227Y, E228Q, or H231Y of SEQ ID NO: 1.

In some embodiments, the modified BoNT polypeptide enters a cell. In some embodiments, the modified BoNT polypeptide does not cleave a SNARE protein.
In some embodiments, the modified BoNT/X polypeptide further comprises a modification in the linker region of (b). In some embodiments, the modification in the linker region comprises one single substitution mutation in a position corresponding to C461 or C467 of SEQ ID NO: 1. In some embodiments, the single substitution mutation corresponds to C461A, C461S, C467A, or C467S in SEQ ID NO: 1. In some embodiments, the modified BoNT/X polypeptide further comprises a modification in the receptor binding domain of (d).

In some embodiments, the modification in the receptor binding domain comprises a substitution mutation in a position corresponding to C1240 of SEQ ID NO: 1.

Other aspects of the present disclosure provide nucleic acid molecules comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%, or 100% identity to the BoNT polypeptide described herein. Nucleic acid vectors comprising such nucleic acid molecules are provided. Cells comprising the nucleic acid molecules or the nucleic acid vectors described herein are provided. In some embodiments, such cells express the BoNT polypeptide described herein.

Methods of producing the BoNT polypeptide of the present disclosure are provided. Such methods comprise the steps of culturing the cell expressing the BoNT polypeptides under conditions wherein said BoNT polypeptide is produced. In some embodiments, the methods further comprise recovering the BoNT polypeptide from the culture.

Further provided herein are use of the modified BoNT polypeptide described herein as a delivery vehicle to deliver therapeutics into neurons.

Some aspects of the present disclosure provide chimeric molecules comprising a first portion linked to a second portion, wherein the first portion is a modified BoNT polypeptide described herein.

In some embodiments, the first portion and the second portion are linked covalently. In some embodiments, the first portion and the second portion are linked non-covalently.

In some embodiments, wherein the second portion is selected from the group consisting of a small molecule, a nucleic acid, a short polypeptide and a protein. In some embodiments, the second portion is a bioactive molecule. In some embodiments, the second portion is a non-polypeptide drug. In some embodiments, the second portion is a therapeutic polypeptide.

Other aspects of the present disclosure provide nucleic acid molecules comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence that has at least
85%, at least 86%, at least 87%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%, or 100% identity to the chimeric BoNT polypeptide described herein. Nucleic acid vectors comprising such nucleic acid molecules are provided. Cells comprising the nucleic acid molecules or the nucleic acid vectors described herein are provided. In some embodiments, such cells express the chimeric BoNT polypeptide described herein.

Methods of producing the chimeric BoNT polypeptide of the present disclosure are provided. Such methods comprise the steps of culturing the cell expressing the chimeric BoNT polypeptides under conditions wherein said chimeric BoNT polypeptide is produced. In some embodiments, the methods further comprise recovering the chimeric BoNT polypeptide from the culture.

Other aspects of the present disclosure provide pharmaceutical compositions comprising the BoNT polypeptides described herein.

In some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

Kit comprising such pharmaceutical compositions and directions for therapeutic administration of the pharmaceutical composition are also provided.

Some aspects of the present disclosure provide methods of treating a condition, comprising administering a therapeutically effective amount of the BoNT polypeptide, the chimeric molecule, or the pharmaceutical composition described herein to a subject to treat the condition.

In some embodiments, the condition is associated with overactive neurons or glands. In some embodiments, the condition is selected from the group consisting of, spasmodic dysphonia, spasmodic torticollis, laryngeal dystonia, oromandibular dysphonia, lingual dystonia, cervical dystonia, focal hand dystonia, blepharospasm, strabismus, hemifacial spasm, eyelid disorder, cerebral palsy, focal spasticity and other voice disorders, spasmodic colitis, neurogenic bladder, anismus, limb spasticity, tics, tremors, bruxism, anal fissure, achalasia, dysphagia and other muscle tone disorders and other disorders characterized by involuntary movements of muscle groups, lacrimation, hyperhydrosis, excessive salivation, excessive gastrointestinal secretions, secretory disorders, pain from muscle spasms, headache pain, dermatological or aesthetic/cosmetic conditions, and obesity/reduced appetite.
In some embodiments, the condition is not associated with unwanted neuronal activity. In some embodiments, the condition is selected from the group consisting of: psoriasis, allergy, haemophagocytic lymphohistiocytosis, and alcoholic pancreatic diseases.

In some embodiments, the administering is via injection to where unwanted neuronal activity is present.

Yet other aspects of the present disclosure provide methods of producing a Clostridial Botulinum neurotoxin (BoNT) polypeptide, the method comprising:

(i) obtaining a first BoNT fragment comprising a light chain (LC) and a N-terminal domain of a heavy chain (H$_N$), wherein the first BoNT fragment comprises a C-terminal LPXTGG (SEQ ID NO: 60) motif;

(ii) obtaining a second BoNT fragment comprising a C-terminal domain of the heavy chain (H$_C$); wherein the second BoNT fragment comprise a specific protease cleavage site at its N-terminus;

(iii) cleaving the second BoNT fragment with a specific protease, wherein the cleavage results in a free Glycine residue at the N-terminus; and

(iv) contacting the first BoNT fragment and the second BoNT fragment in the presence of a transpeptidase, thereby ligating the first BoNT fragment and the second BoNT fragment to form a ligated BoNT.

In some embodiments, the first BoNT fragment further comprises an affinity tag. In some embodiments, the affinity tag is fused to the first BoNT fragment at the N-terminus. In some embodiments, the affinity tag is fused to the first BoNT fragment at the C-terminus. In some embodiments, the affinity tag is selected from the group consisting of: His6, GST, Avi, Strep, S, MBP, Sumo, FLAG, HA, Myc, SBP, E, Calmodulin, Softag 1, Softag 3, TC, V5, VSV, Xpress, Halo, and Fc.

In some embodiments, the second BoNT fragment further comprises an affinity tag. In some embodiments, the affinity tag is fused to the first BoNT fragment at the N-terminus. In some embodiments, the affinity tag is fused to the second BoNT fragment at the C-terminus. In some embodiments, the affinity tag is selected from the group consisting of: His6, GST, Avi, Strep, S, MBP, Sumo, FLAG, HA, Myc, SBP, E, Calmodulin, Softag 1, Softag 3, TC, V5, VSV, Xpress, Halo, and Fc.

In some embodiments, the protease is selected from the group consisting of: thrombin, TEV, PreScission, MMP-12, MMP-13, MMP-17, MMP-20, Granzyme-B, Enterokinase, and SUMO protease. In some embodiments, the cognate protease is thrombin.
In some embodiments, the first BoNT fragment is from BoNT serotype A, B, C, D, E, F, G, or X. In some embodiments, the first BoNT fragment is from BoNT/X. In some embodiments, the second BoNT fragment is from BoNT serotype A, B, C, D, E, F, G, or X. In some embodiments, the second BoNT fragment is from BoNT/A. In some embodiments, the second BoNT fragment is from BoNT/B. In some embodiments, the second BoNT fragment is from BoNT/C. In some embodiments, the second BoNT fragment is from BoNT/X.

In some embodiments, the transpeptidase is a sortase. In some embodiments, the sortase is from Staphylococcus aureus (SrtA).

These and other aspects of the disclosure, as well as various advantages and utilities will be apparent with reference to the Detailed Description of the Invention. Each aspect of the disclosure can encompass various embodiments as will be understood.

All documents identified in this application are incorporated in their entirety herein by reference.

**BRIEF DESCRIPTION OF THE DRAWINGS**

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure, which can be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein. The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

**FIGS.** 1A-1E show the identification of BoNT/X as a new BoNT. **FIG.** 1A shows a phylogenic tree of the protein sequence alignment for BoNT/A-G, BoNT/F5, TeNT, and BoNT/X, analyzed by ClustalW method. The percentages of sequence identity between each toxin and BoNT/X are denoted after each toxin. The percentages of sequence identity between BoNT/E and BoNT/F, and between BoNT/B and BoNT/G were also noted. **FIG.** IB, upper panel, shows a schematic drawing of the three domains of BoNT/X, with conserved protease motif in the LC and the ganglioside binding motif in the He noted. **FIG.** IB, lower panel, shows a sliding sequence comparison window demonstrating that BoNT/X has a low similarity evenly distributed along its sequence to all other seven BoNTs and TeNT. **FIG.** 1C is a schematic drawing of the orf gene cluster that hosts BoNT/X gene (upper panel), which has two unique features compared to two known variants of orfX cluster (middle and lower panels): (1) there is an additional orfX2 protein (designated as orfX2b) located next to the
BoNT/X gene; (2) the reading frame of orfX genes has the same direction with BoNT/X gene. FIG. 1D is a schematic illustrating the unique gene directionality and additional OrfX2 gene found in BoNT/X. FIG. 1E shows a preliminary structure of the BoNT/X light chain. The dark dot represents the active site zinc. The structure is shown at a 1.9Å resolution.

FIGS. 2A-2J show the LC of BoNT/X (X-LC) cleaves VAMPs at a unique site. FIG. 2A shows X-LC, with or without pre-treated with EDTA, incubated with rat brain detergent extracts (BDE). Immunoblot analysis was carried out to detect syntaxin 1, SNAP-25, and VAMP2. Synaptophysin (Syp) was also detected as a loading control. The LC of BoNT/A (A-LC) and BoNT/B (B-LC) were analyzed in parallel. Cleavage of VAMP2 by B-LC results in loss of immunoblot signals, while cleavage of SNAP-25 by A-LC generates a smaller fragment of SNAP-25 that can still be detected on immunoblot (marked by an asterisk). Incubation with X-LC resulted in loss of VAMP2 immunoblot signals, suggesting that X-LC cleaved VAMP2. EDTA blocked the activity of X-, A-, and B-LCs. FIG. 2B shows VAMP2 (residues 1-96) purified as a His6-tagged recombinant protein and incubated with X-LC. Samples were analyzed by SDS-PAGE and Coomassie Blue staining. X-LC converted VAMP2 (1-96) into two smaller fragments, indicating that X-LC cleaved VAMP2. FIGs. 2C-2E show VAMP2 (1-96) incubated with X-LC. Whole protein samples were then analyzed by mass spectrometry (LC-MS/MS) to determine the precise molecular weight of cleaved fragments. Eluted peptide peaks from the HPLC column were plotted in FIG. 2C over running time (RT, X-axis). The mass spectrometry data for the two cleavage products are shown in FIGs. 2D and 2E, respectively, with mass-to-charge ratio (m/z) noted for each signal. The molecular weight is deducted by multiplying m with z, followed by subtracting z. The protein sequences for the two cleavage products correspond to SEQ ID NO: 61 and 62 from top to bottom and are shown in FIG. 2C. FIG. 2F shows a sequence alignment between VAMP 1, 2, 3, 4, 5, 7, 8, Sec22b and Ykt6, with cleavage sites for BoNT/B, D, F, G, and X underlined, and two SNARE motifs boxed. The sequences correspond to SEQ ID NOs: 63-71 from top to bottom. FIG. 2G shows HA-tagged VAMP1, 3, 7, and 8, and myc-tagged Sec22b and Ykt6 expressed in HEK293 cells via transient transfection. Cell lysates were incubated with X-LC and subjected to immunoblot analysis detecting the HA or Myc tag. Actin served as a loading control. X-LC cleaved VAMP1, 3 and Ykt6, but not VAMP7, 8 and Sec22. FIG. 2H shows GST-tagged Ykt6 incubated with X-LC (100 nM) for the indicated times. Samples were analyzed by SDS-PAGE and Coomassie Blue staining. X-LC cleaved Ykt6. FIG. 2I shows GST-tagged cytoplasmic domains of VAMP2 (33-86), VAMP4 (1-115), and VAMP5 (1-70) incubated with X-LC for
the indicated times. Samples were analyzed by SDS-PAGE and Coomassie Blue staining. X-LC cleaved both VAMP4 and VAMP5. A longer incubation time (360 min) is required to cleave majority of VAMP5. Note that VAMP5 protein contains an additional band that is either degradation product or bacterial protein contaminant, which runs close (but not identical) to the cleavage product on SDS-PAGE. FIG. 2J shows experiments carried out as described in FIG. 2A, except that VAMP4 and Sec22b were detected. Synaptotagmin I (Syt I) was detected as a loading control. X-LC cleaved native VAMP4 in BDE.

FIGs. 3A-3E show activation of BoNT/X by proteolytic cleavage of the linker region between LC and H_N. FIG. 3A shows a sequence alignment of the linker regions between LC and H_N of the seven BoNTs and BoNT/X. The sequences correspond to SEQ IDs NOs: 72-79 from top to bottom. BoNT/X has the longest linker region among all BoNTs, which contains an extra cysteine in addition to the two conserved cysteines in the LC and in the H_N. The Lys-C cutting site under limited proteolysis was identified by mass spectrometry approach. FIG. 3B shows cultured rat cortical neurons exposed to indicated concentrations of X-LC-H_N in media for 12 hours. Cell lysates were harvested and immunoblot analysis was carried out to examine syntaxin 1, SNAP-25, and VAMP2 in neurons. Actin served as a loading control. Trypsin-activated LC-H_N of BoNT/A (A-LC-H_N) and BoNT/B (B-LC-H_N) were analyzed in parallel as controls. X-LC-H_N entered neurons and cleaved VAMP2, as evidenced by loss of VAMP2 immunoblot signals. X-LC-H_N activated by Lys-C showed a drastically increased potency than non-activated X-LC-H_N. X-LC-H_N is more potent than trypsin-activated B-LC-H_N and A-LC-H_N, which did not show any detectable cleavage of their SNARE substrates in neurons under the same assay concentrations. FIG. 3C shows X-LC-H_N mutants with indicated cysteine mutated, as well as WT X-LC-H_N, activated by limited proteolysis and analyzed by SDS-PAGE and Coomassie Blue staining, with or without DTT. C423S mutation resulted in two 50 kDa fragments, with or without DTT. Mutants harboring C461S or C467S showed a single band at 100 kDa in the absence of DTT, and it separated into two 50 kDa bands in the presence of DTT, demonstrating that both C461 and C467 on the H_N can form the inter-chain disulfide bond with C423 on the LC. A portion of WT X-LC-H_N formed aggregates at the top of the SDS-PAGE gel. These aggregates are due to formation of inter-molecular disulfide bond, as they disappeared in the presence of DTT. Mutating any one of three cysteines abolished aggregates, indicating that formation of inter-molecular disulfide bond is due to existence of an extra cysteine in the linker region. The majority of activated WT X-LC-H_N also separated to two 50 kDa bands on SDS-PAGE gel without DTT. This is due to disulfide
bond shuffling described in FIG. 3D. FIG. 3D shows WT X-LC-H_N activated by limited proteolysis, followed by pre-incubation with indicated concentrations of NEM to block disulfide bond shuffling. The samples were then analyzed by SDS-PAGE and Coomassie Blue staining, with or without the presence of DTT. Majority of WT X-LC-H_N exist as a single band at 100 kDa without DTT after NEM treatment, indicating that WT X-LC-H_N mainly contains inter-chain disulfide bond. FIG. 3E shows experiments carried out as described in FIG. 3B, except that neurons were exposed to either WT or indicated X-LC-H_N mutants. Mutating the cysteine on the LC (C423) abolished the activity of X-LC-H_N, while mutating one of the two cysteines on the H_N (C461 or C467) did not affect the activity of X-LC-H_N on neurons. These results confirmed that formation of the inter-chain disulfide bond is essential for the activity of X-LC-H_N.

FIGs. 4A-4F show full-length BoNT/X is active on cultured neurons and in vivo in mice. The sequences are as follows: LVPR-GS (SEQ ID NO: 80), LPETGG-His6 (SEQ ID NO: 81), GG-His6 (SEQ ID NO: 82) and LPETGS (SEQ ID NO: 59). FIG. 4A shows a schematic drawing illustrates synthesis of full-length BoNT/X using sortase ligation method. FIG. 4B shows that sortase ligation reaction mixture and indicated control components were analyzed by SDS-PAGE and Coomassie Blue staining. The asterisk marks aggregates of proteins due to inter-molecular disulfide bond, as these aggregates disappeared in the presence of DTT. The molecular weight marker is in lane 1 (starting from the left side). Full-length BoNT/X (X-FL) only appeared in the sortase ligation mixture (lane 7 and lane 14). FIG. 4C shows that neurons exposed to the same amount (5 µl) of sortase ligation mixture or indicated control components for 12 hours in media. Cell lysates were analyzed by immunoblot. X-LC-HN alone cleaved some VAMP2 due to its high concentration in the reaction mixture. The control mixture containing both X-LC-HN and X-HC but not sortase, slightly enhanced cleavage of VAMP2 as compared to X-LC-HN alone, likely because X-HC associates with X-LC-HN via non-covalent interactions. Ligating X-LC-HN and X-HC by sortase enhanced cleavage of VAMP2 over the mixture of X-LC-HN and X-HC without sortase, demonstrating that ligated X-FL is functional in neurons. FIG. 4D shows that sortase reaction mixture as prepared as described in panel b (lane 7) is active in vivo analyzed using DAS assay in mice. The injected limb developed flaccid paralysis and the toes failed to spread within 12 hours. The left limb was not injected with toxins, serving as a control. FIG. 4E shows that BoNT/A-G, a mosaic toxin BoNT/DC, and BoNT/X were subjected to dot blot analysis (0.2 µg per toxin, spotted on nitrocellulose membranes), using four horse antisera (trivalent anti-BoNT/A, B, and
E, anti-BoNT/C, anti-BoNT/DC, and anti-BoNT/F), as well as two goat antisera (anti-BoNT/G and anti-BoNT/D). BoNT/X is composed of purified X-LC-HN and X-HC at 1:1 ratio. These antisera recognized their corresponding target toxins, yet none of them recognized BoNT/X. FIG. 4F shows that full-length inactive form of BoNT/X (BoNT/XRY) was purified as a His6-tagged recombinant protein in E.coli and analyzed by SDS-PAGE and Coomassie Blue staining, with or without DTT.

FIG. 5 is a phylogenetic tree showing the distribution and relationship of Clostridial neurotoxins. The tree represents the relationships of different BoNTs and TeNT sequences from the Jackhmmer search. BoNT/X is circled.

FIG. 6 shows a mass spectrometry analysis of intact VAMP2 (1-96). His6-tagged VAMP2 (1-96) was analyzed by LC-MS/MS mass spectrometry. The HPLC profile is listed in the left panel, together with the protein sequence. The mass spectrometry data for full-length VAMP2 (1-96) was shown in the right panel and corresponds to SEQ ID NO: 83, with m/z value marked for each signal.

FIGs. 7A-7F shows the identification of the cleavage site on GST-VAMP2 (33-86) by X-LC. FIG. 7A shows a GST-tagged VAMP2 (33-86) incubated with or without X-LC. Samples were analyzed by SDS-PAGE and Coomassie Blue staining. FIGs. 7B-7C show intact GST-tagged VAMP2 (33-86) analyzed by LC-MS/MS mass spectrometry. The HPLC profile was shown in FIG. 7B. The mass spectrometry data was shown in FIG. 7C, with protein sequence (SEQ ID NO: 84) noted in FIG. 7C. VAMP2 (33-66) and VAMP2 (67-86) are marked. FIGs. 7D-7E show GST-tagged VAMP2 (33-86) incubated with X-LC. Samples were then analyzed by LC-MS/MS mass spectrometry. The HPLC profile is shown in FIG. 7D. The mass spectrometry data for the C-terminal fragment (SEQ ID NO: 85) generated by X-LC is shown in FIG. 7E. The mass spectrometry data for the N-terminal fragment (SEQ ID NO: 86) was shown in FIG. 7F. The protein sequences of the C- and N-terminal fragments were indicated in FIGs. 7E-7F, and correspond to SEQ ID NOs: 85 and 86 respectively.

FIGs. 8A-8B show that XA chimeric toxin is active on neurons. FIG. 8A shows a XA chimeric toxin generated by ligating X-LC-HN with A-Hc by sortase, similar to generating X-FL as described in FIG. 4A. The sortase ligation mixture and indicated control components were analyzed by SDS-PAGE and Coomassie Blue staining. The ligation is efficient as majority of X-LC-HN was ligated into XA chimeric toxin. FIG. 8B shows rat cortical neurons exposed to the indicated control components or sortase ligated XA mixture (5 µl) for 12 hours
in media. Cell lysates were analyzed by immunoblot. X-LC-H<sub>N</sub> alone cleaved some VAMP2 due to its high concentration in the reaction mixture. Ligated XA cleaved VAMP2 in neurons.

FIG. 9 shows that mutating the extra cysteine in the H<sub>N</sub> and the cysteine in the He does not affect activity of BoNT/X. X-H<sub>c</sub> (C1240S) was ligated with WT X-LC-H<sub>N</sub>, X-LC-H<sub>N</sub> (C461S), or X-LC-H<sub>N</sub> (C467S) by sortase ligation. Neurons were exposed to sortase ligation mixture or control components (5 µM) for 12 hours in media. Cell lysates were analyzed by immunoblot. Mutating C1240 and one of the cysteine on H<sub>N</sub> (C461 or C467) did not affect the activity of BoNT/X, as ligated mutant toxins are capable of entering neurons and cleaved VAMP2.

FIG. 10 shows antisera raised against the seven serotypes of BoNTs neutralizing their target BoNTs on neurons. Cultured rat cortical neurons were exposed to indicated BoNTs, with or without pre-incubation with indicated antisera. Cell lysates were harvested 12 hours later and subjected to immunoblot analysis. All antisera specifically neutralized their target BoNTs, without affecting the activity of a different serotype of BoNTs, thus validating the specificity and potency of these antisera. The concentrations for BoNTs were: BoNT/A (50 pM), BoNT/B (2 nM), BoNT/C (1.5 nM), BoNT/D (100 pM), BoNT/E (0.5 nM), BoNT/F (0.5 nM), BoNT/G (5 nM). The antiserum against BoNT/A/B/E was used at 20 µM per well. All the other antisera were used at 10 µM per well. BoNTs were pre-incubated with indicated antisera for 30 mins at 37 °C prior to adding into culture media.

FIGs. 11A-1 IC show that BONT/XRY is not active on neurons. FIG. 11A shows cultured rat cortical neurons exposed to BONT/XRY at indicated concentrations. Cell lysates were analyzed by immunoblot. VAMP2 was not cleaved, indicating that BoNT/X<sub>RY</sub> is not active on neurons. FIG. 11B shows the SDS-PAGE analysis of cell lysate and supernatant (S/N) expression of BONT/XRY (4-12% BisTris, MOPS buffer). A band at 150kDa corresponding to BoNT/X is clearly visible in both lysate and soluble fraction. FIG. 11C shows the SDS-PAGE analysis of a final sample of highly purified BoNT/X<sub>RY</sub> (4-12% BisTris, MOPS buffer). A single band at 150 kDa corresponding to BoNT/X is clearly visible and shows -90% purity.

FIGs. 12A-12F show that BoNT/X binds to all four brain gangliosides. FIGs. 12A-12D show BoNT/X (squares), and A-Hc (circles) binding to GDla (FIG. 12A), GTlb (FIG. 12B), GDib (FIG. 12C), and GM1 (FIG. 12D), respectively. Curves correspond to an average of triplicate ELISA assays and were fitted with Prism7 (GraphPad software). FIG 12E shows a
summary of BoNT/X binding to all four gangliosides compared with the overall binding of BoNT/A in FIG 12F.

FIGs. 13A-13D show the identification of the cleavage sites of X-LC on Ykt6 by mass spectrometry analysis. FIGs. 13A-13D show 10 µg GST-tagged Ykt6 (1-192), with or without pre-incubation with X-LC, were separated on SDS-PAGE (FIG. 13A). The protein bands were excised as indicated and digested by chymotrypsin. Digested peptides were desalted and analyzed by reversed phase HPLC via C18 column coupled with ESI-MS. The HPLC profiles of GST-Ykt6 without pre-treatment with X-LC was shown in FIG. 13B, and the sample pretreated with X-LC was shown in FIG. 13C. One peptide was identified to be 100-fold higher intensity in the samples pre-treated with X-LC than in the samples that was not exposed to X-LC (denoted with an asterisk). This peptide was eluted at 37 min RT, with m/z = 611 (FIG. 13D), which can only fit the peptide sequence ESLERGEKLDDLVSK (SEQ ID NO: 87) in Ykt6, indicating that this is the peptide located at the N-terminal side of the cleavage site for X-LC. Therefore the cleavage site is K173-S174 in Ykt6.

FIGs. 14A-14E show the identification of the cleavage sites of X-LC on VAMP4 and VAMP5 by mass spectrometry analysis. FIGs. 14A-14E show experiments carried out as described in FIG. 13, except that VAMP4 (FIGs. 14B, 14C) and VAMP5 (FIGs. 14D, 14E) were analyzed. FIG. 14B is the peptide that marks the N-terminal site of the cleavage site in VAMP4. The sequence of the peptide DELQDK corresponds to SEQ ID NO: 88. FIG. 14C is the peptide that marks the C-terminal site of the cleavage site in VAMP5. The sequence of the peptide SESLSDNATAF corresponds to SEQ ID NO: 89. FIG. 14D is the peptide that marks the N-terminal site of the cleavage site in VAMP5. The sequence of the peptide AELQQR corresponds to SEQ ID NO: 90. FIG. 14E is the peptide that marks the C-terminal site of the cleavage site in VAMP5. The sequence of the peptide SDQLLDMSTF corresponds to SEQ ID NO: 91. Thus, the cleavage sites were determined to be K87-S88 in VAMP4 and R40-S41 in VAMP5.

DETAILED DESCRIPTION OF SOME EMBODIMENTS

*Clostridium Botulinum* neurotoxins (BoNTs) are a family of bacterial toxins produced by *Clostridium* bacteria, with seven well-established serotypes (BoNT/A-G) \(^1-3\). They are one of the most dangerous potential bio-terrorism agents, classified as a "Category A" select agent by Center for Disease Control (CDC) of United States \(^4\). These toxins are produced as a single polypeptide and can be separated by bacterial or host proteases into a light chain (LC, ~ 50
kDa) and a heavy chain (He, ~100 kDa). The two chains remain connected via an inter-chain disulfide bond. The He contains two sub-domains: the N-terminal $H_N$ domain that mediates translocation of the LC across endosomal membranes, and the C-terminal He domain that mediates binding to receptors on neurons. The inter-chain disulfide bond is reduced once the LC translocates into the cytosol.\(^5\)\(^6\). Released LC acts as a protease to specifically cleave a set of neuronal proteins: BoNT/A, C, and E cleave at distinct sites on a protein known as SNAP-25; BoNT/B, D, F, and G cleave at different sites on a vesicle protein VAMP; and BoNT/C also cleaves a transmembrane protein syntaxin 1.\(^1\)\(^-\)\(^3\). These three proteins form a complex, known as SNARE complex, which is essential for release of neurotransmitters.\(^7\)\(^8\). Cleavage of any one of these three SNARE proteins blocks neurotransmitters release from neurons, thus paralyzing muscles.

BoNTs are the most potent toxins known and cause the human and animal disease known as botulism.\(^3\) The major form of botulism is caused by ingesting food contaminated with BoNTs (food botulism). Other forms also exist such as infant botulism, which is due to colonization of the intestine by toxin-producing bacteria in infants. BoNTs are always produced together with another 150 kDa protein known as NTNHA (non-toxic non-hemagglutinin protein), which forms a pH-dependent complex with BoNTs and protects BoNTs from proteases in the gastrointestinal tract.\(^9\) Genes encoding BoNT and NTNHA are found in two types of gene clusters: (1) HA cluster, containing genes for three conserved proteins HA17, HA33 and HA70, which form a complex with BoNT/NTNHA and facilitate absorption of toxins across the intestinal epithelial barrier.\(^10\)\(^-\)\(^12\). (2) OrfX cluster, which encodes conserved OrfX1, OrfX2, OrfX3 and P47 proteins with unknown functions.\(^13\)

Because local injections of minute amounts of toxins can attenuate neuronal activity in targeted regions, BoNTs have been used to treat a growing list of medical conditions, including muscle spasms, chronic pain, overactive bladder problems, as well as for cosmetic applications. The market for BoNTs has already surpassed $1.5 billion in 2011 and is projected to reach 2.9 billion by 2018.

BoNTs were traditionally typed by neutralization assays in mice, by injecting culture supernatant of Clostridium bacteria into mice, with or without antisera against known BoNTs. The first distinguished serotypes, BoNT/A and BoNT/B, were established in 1919 by Georgina Burke.\(^18\) The last of the seven type, BoNT/G, was recognized in 1969 from soil samples in Argentina.\(^19\) No new serotype of BoNTs has been recognized since 1970. This classification held true after protein sequences for each BoNT was determined in 1990's. The sequence
identity between any two pairs among the seven BoNTs ranges from 32% to 65.3%. All seven BoNTs have been identified and characterized before the era of their medical use. Therefore, there is no patent on any of these toxins. Any company is free to produce and market any one of these seven BoNTs. Among the seven types, BoNT/A and BoNT/B are the two toxins that are currently FDA-approved for use in humans. BoNT/A is the dominant type used for both medical and cosmetic applications, marketed as Botox from Allergan Inc., Dysport from IPSEN Inc., and Xeomin from Merz Inc. BoNT/B is marketed as Myobloc by USWorld Med. There are considerable interests in developing other BoNT types as therapeutic toxins, for two major reasons:

1. A major limitation in treatment is generation of neutralizing antibody against BoNT/A or BoNT/B in patients, which renders future treatment with the same toxin ineffective. In this case, patients will need to be treated with a different type of BoNTs. This is why BoNT/B is often utilized to treat patients who have generated neutralizing antibodies against BoNT/A during treatment, but there is a need for alternative toxins for patients who have generated antibodies against both BoNT/A and BoNT/B.

2. Although all BoNTs share the same structure and function, there are also considerable differences between them. For instance, BoNT/A cleaves SNAP-25 and uses a protein SV2 as its receptor, whereas BoNT/B cleaves VAMP and uses a protein synaptotagmin (Syt) as its receptor. These functional variations may translate to potential differences in therapeutic efficacy targeting distinct types of neurons. In addition, the stability and therapeutic duration can be also different among seven types of toxins. Therefore, a different toxin type may have its advantage over BoNT/A and BoNT/B.

Rapid progress on genomic sequencing in recent years has revealed a remarkable diversity of BoNTs. First, there are multiple subtypes, which can be recognized by the same antiserum, but contain significant levels of variations on protein sequences (2.6%~31.6% differences). For instance, BoNT/A contains 8 known subtypes, designated as BoNT/A1-A8. Furthermore, multiple mosaic toxins exist, likely derived from recombination of toxin genes. For instance, a "type H" was reported in 2013, but it was later recognized as a chimeric toxin because its LC shares ~ 80% identity with the LC of a BoNT/F subtype, BoNT/F5, and its Hc shares ~ 84% identity to the Hc of BoNT/Al. Consistently, this toxin can be recognized and neutralized by available antisera against BoNT/A.

The gene cluster encoding BoNTs can be on plasmids, bacterial phage, or chromosomes, indicating that the toxin genes are mobile and subject to horizontal gene transfer.
There are also cases that a Clostridium bacteria strain contains two or even three different toxin genes \(^{32,35,36}\). In these cases, one toxin is usually expressed at higher levels (designed with a capital letter) than the other toxin (designated with a lower case letter). For instance, strains that express high levels of BoNT/B and low levels of BoNT/F are known as BoNT/Bf strains. There are also cases that one toxin is expressed, but the other toxin is not expressed, which is known as silent toxin (usually marked with \((\))\). For instance, a survey for infant botulism cases in California showed that 8% strains were BoNT/A(B), which means these strains contain genes for both BoNT/A and BoNT/B, but only express detectable levels of BoNT/A \(^{37-39}\).

As illustrated in the drawings and examples of the present disclosure, published Clostridium bacteria genomic sequence databases were searched, and a novel BoNT gene (hereafter designated "BoNT/X") encoded on the chromosome of *Clostridium botulinum* strain 111 was identified. Strain 111 was first isolated from an infant botulism patient in Japan in 1996 \(^{40}\). It has been shown that toxicity from strain 111 in mice can be neutralized by BoNT/B antisera \(^{40}\). It was later confirmed that this strain expresses a subtype of BoNT/B, BoNT/B2, encoded on a plasmid \(^{41,42}\). The sequence of BoNT/X was deposited into PubMed database in February of 2015, as part of genomic sequence of Strain 111. BoNT/X has not been characterized before. It remains unknown whether it is expressed in the strain 111 and whether it is a functional toxin.

Also provided herein are the characterization of BoNT/X at functional levels. Its LC was found to cleave VAMP at a site distinct from known target sites of all other BoNTs. The full-length toxin, produced by covalently linking non-toxic fragments via sortase, was found to enter cultured neurons and cleave VAMP in neurons, inducing flaccid paralysis in mice. Finally, it was found that the toxin is not recognized by antisera raised against all seven known BoNTs, establishing BoNT/X as a novel BoNT serotype. Its identification poses an urgent challenge for developing effective countermeasures. It also has the potential to be developed into a new therapeutic toxin and can be used to generate chimeric toxins with potentially distinct pharmacological properties.

As used herein, the term "Clostridial Botulinum neurotoxin (BoNT) polypeptide" encompasses any polypeptide or fragment from a Botulinum neurotoxin described herein. In some embodiments, the term BoNT refers to a full-length BoNT. In some embodiments, the term BoNT refers to a fragment of the BoNT that can execute the overall cellular mechanism whereby a BoNT enters a neuron and inhibits neurotransmitter release. In some embodiments, the term BoNT simply refers to a fragment of the BoNT, without requiring the fragment to
have any specific function or activity. For example, a BoNT polypeptide may refer to the light chain (LC) of a BoNT, e.g., BoNT/X. Other terms that may be used throughout the present disclosure for "Clostridial Botulinum neurotoxins" may be BoNTs, Botulinum toxins, or C. Botulium toxins. It is to be understood that these terms are used interchangeably. "BoNT/X" refers to the novel BoNT serotype described and characterized in the present disclosure. The BoNT/X protein sequence (GenBank No. BAQ12790.1; four cysteines are underlined and bolded) is also provided:

```
MKLIEINKFNYNDPIDDGINVITMRPPRHSKDINGKGGPFKAFOQIVKNIIVIPERYYNFTNTNDDLNIPSEPME
ADAIYNPNYLNTPEKDEFLQGVIKVLERIKSKPEGKLLLEISSSLIPPLVSNGLATLSDNEITAYQENNNI
VSNLQANLVIYGPGDPDIAANNATGYLSTPIISNNEGTLSEVFSYPSFYLKPFDESGYNYRSLVNLNNVFKVRE
FAPDPASTMHELHVTHNLGYISRNFYYFNFTGTGIESTRQRQNSLIFEELTTFGGIDSKAISLLHKKIET
AKNNYYTLISERLNTVTENDDLLKYIKNKIPVQGRLGNFKLDTAEFEKKNLNTILFVLNESNLQRFSILVR
KHYKLERPIDPIYYVLIDDNSYSTLEGFNSSQGNSDFQGQLLESSFYFEKIESNALRAFICKPNNGLLLYNAI
YRNSKNYLNNIDLEDKKTSTKTNVYPSILLNGIDEVENKDLFLISNKDLNDINLSEEKIKPETTYVFKKD
KLPPQDITLSNFDEASNSIPSISQONLISNEELNLEPIRSNSLFEIKITYVDKLTPPTFTERAQIDSSSISR
VETLSDVDEALSNNPKVYSPKMNMSNTINSIETGITSTYIFYYQWLRSLIKDFDSTRGTGIDVIDKSDSLTALVP
YIGPLLNIANDIRHGDFVGAIELAGITALLEYPETFSTLPGLEVIGGELAREQVEIAVNNALDKRDQKW
AAEVNITQAWGGTHILQINTTRLAHTYKALSRQANAIKMNMFQLANYKGNNIDKAKKAIKNAISETELL
KSVEQAMKNTKFMKLSNYLTKEMIPKQDNLKNFLETTKTLDFKEKEDILGTNLSSSLRRLKVR
FLKIVNIAFDINIDPFPSEDDLINGQKNEIEDYEVNMLGAEDGKIKDLGTSOIEWGSDLADIADGRENKAIIK
GENSTIJKIAMNKLRELSATDNSFSIFWKHPKTNLLNNGEYTVELENFQGRWKIIQDSKLIWYLRDH
NNSIKIVTPDYIAFWGNWLITITTNRSSKGSIYVNGSKIEKDISSWTVEDDDIPFRLKNRNRDTQAITLD
QFSIYRKEQNQNEVVKLYNYFYNSYIRDJWQNFLQYNNKKYLYLTQDCKPGKLGREYWSFGYDYL
DSKTIFFFFNWYNGSVKLKSNKSKLDGLVRFNQKDFQLEIDGYNMGIASCDFRNEDTYGITTGH
DLTTDEIFFQREKYRNCYQKTPYINIFHKSMLYMSTETSKPTFHDYRDWVYSSAWFQNYENKLKHT
KTNWYFIPKDEGWED (SEQ ID NO: 1)
```

A "modified Clostridial Botulinum neurotoxin (BoNT)" encompasses a BoNT comprising any modifications in the amino acid sequence, e.g., truncation, addition, amino acid substitution, and any combination thereof. For example, a BoNT/X comprising amino acid substitution mutations in C461 or C467 is a modified BoNT. In another example, a fragment or a domain of the full-length BoNT (e.g., the protease, or LC) is considered a modified BoNT. In some embodiments, at least one BoNT may also comprise amino acid substitution mutations, e.g., a protease domain comprising substitution mutations at positions C461 or C467 of BoNT.

The term "enters a cell" when used to describe the action of a BoNT of the present
disclosure, encompasses the binding of a BoNT to a low or high affinity receptor complex, binding of a BoNT to ganglioside, the internalization of the toxin, the translocation of the toxin light chain into the cytoplasm and the enzymatic modification of a BoNT substrate.

As used herein, the term "Clostridial Botulinum neurotoxin (BoNT) protease domain" is synonymous to "light-chain (LC)." The BoNT protease domain is located in the light chain of the BoNT, and thus is also referred to as the LC. The term means a BoNT domain that can execute the enzymatic target modification step of the intoxication process. If the LC from a specific BoNT serotype is referred to, the term "serotype-LC" is used. For example, "X-LC" means the LC polypeptide from BoNT/X. A BoNT protease domain specifically targets a C. Botulinum toxin substrate and encompasses the proteolytic cleavage of a C. Botulinum toxin substrate, such as, e.g., SNARE proteins such as a SNAP-25 substrate, a VAMP substrate and a Syntaxin substrate. In BoNT (e.g., BoNT/X, BoNT/A, BoNT/B, BoNT/C, etc.). The protease domain or the LC is considered to correspond to about amino acid 1-439 of BoNT/X. The domain boundary may vary by about 25 amino acids. For example, the protease domain may correspond to amino acids 1-414 or 1-464 of BoNT/X. In some embodiments, the protease domain may correspond to amino acids 1-438, 1-437, 1-436, 1-435, 1-434, 1-433, 1-432, 1-431, 1-430, 1-429, 1-439, 1-440, 1-441, 1-442, 1-443, 1-444, 1-445, 1-446, 1-447, 1-448, or 1-449 of BoNT/X.

As used herein, the term "Clostridial Botulinum neurotoxin (BoNT) translocation domain" is synonymous with "H_N domain" and means a BoNT domain that can execute the translocation step of the intoxication process that mediates BoNT light chain translocation. Thus, an H_N facilitates the movement of a BoNT light chain across a membrane into the cytoplasm of a cell. Non-limiting examples of a H_N include a BoNT/A H_N, a BoNT/B H_N, a BoNT/C1 H_N, a BoNT/D H_N, a BoNT/E H_N, a BoNT/F H_N, a BoNT/G H_N, and a BoNT/X H_N. The translocation domain is located in the N-terminus of the heavy chain (He), and thus is also referred as H_N. It is to be understood that these terms are used interchangeably herein.

As used herein, the term "linker region" refers to the amino acid sequence between the BoNT protease domain and the translocation domain. The linker comprises two cysteines at position 461 and 467, one of which forms an inter-molecular disulfide bond with a cysteine in the protease domain, C423 (C461-C423 disulfide bond, or C467-C423 disulfide bond). The formation of this disulfide bond is essential for the activity of BoNT/X.

As used herein, the term "LC-HN" refers to a BoNT polypeptide encompassing the protease domain, the linker region, and the translocation domain. If the LC-HN from a specific
BoNT serotype is referred to, the term "serotype-LC-H \(N\)" is used. For example, "X-LC-H \(N\)" means the LC-\(H_N\) polypeptide from BoNT/X. The LC-\(H_N\) polypeptide is considered to correspond to about amino acid 1-892 of BoNT/X. The domain boundary may vary by about 25 amino acids. For example, LC-HN polypeptide may correspond to about amino acid 1-917 or 1-867 of BoNT/X. In some embodiments, the LC-\(H_N\) polypeptide may correspond to amino acids 1-893, 1-894, 1-895, 1-896, 1-897, 1-898, 1-899, 1-900, 1-901, 1-892, 1-891, 1-890, 1-889, 1-888, 1-887, 1-886, 1-885, 1-884, or 1-883 of BoNT/X.

As used herein, the term "Clostridial Botulinum neurotoxin (BoNT) receptor-binding domain" is synonymous with "He domain" and means any naturally occurring BoNT receptor binding domain that can execute the cell binding step of the intoxication process, including, e.g., the binding of the BoNT to a BoNT-specific receptor system located on the plasma membrane surface of a target cell. Some aspects of present disclosure relate to modified BoNT receptor binding domains from serotype X (BoNT/X). In some embodiments, a "modified BoNT/X receptor binding domain" comprises amino acid substitutions in a position corresponding to C1240 in BoNT/X (SEQ ID NO: 1). The receptor binding domain, or the \(\frac{3}{4}\), is considered to correspond to about amino acid 893-1306 of BoNT/X. The domain boundary may vary by about 25 amino acids. For example, the receptor binding domain or He may correspond to amino acids 868-1306 or 918-1306. In some embodiments, the receptor binding domain or Hc may correspond to amino acids 893-1306, 894-1306, 895-1306, 896-1306, 897-1306, 898-1306, 899-1306, 900-1306, 901-1306, 902-1306, 892-1306, 891-1306, 890-1306, 889-1306, 888-1306, 887-1306, 886-1306, 885-1306, 884-1306, or 883-1306 of BoNT/X.

By "isolated" is meant a material that is free to varying degrees from components which normally accompany it as found in its native state. "Isolate" denotes a degree of separation from original source or surroundings, e.g., from a cell or from flanking DNA or from the natural source of the DNA. The term "purified" is used to refer to a substance such as a polypeptide that is "substantially pure", with respect to other components of a preparation (e.g., other polypeptides). It can refer to a polypeptide that is at least about 50%, 60%, 70%, or 75%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% pure, with respect to other components. The terms "substantially pure" or "essentially purified", with regard to a polypeptide, refers to a preparation that contains fewer than about 20%, more preferably fewer than about 15%, 10%, 8%, 7%, most preferably fewer than about 5%, 4%, 3%, 2%, 1%, or less than 1%, of one or more other components (e.g., other polypeptides or cellular components).
The term "substitution mutation" without the reference to a specific amino acid, may include any amino acid other than the wild type residue normally found at that position. Such substitutions may be replacement with non-polar (hydrophobic) amino acids, such as glycine, alanine, valine, leucine, isoleucine, methionine, phenylalanine, tryptophan, and proline.

Substitutions may be replacement with polar (hydrophilic) amino acids such as serine, threonine, cysteine, tyrosine, asparagine, and glutamine. Substitutions may be replacement with electrically charged amino acids, e.g., negatively electrically charged amino acids such as aspartic acid and glutamic acid and positively electrically charged amino acids such as lysine, arginine, and histidine.

The substitution mutations described herein will typically be replacement with a different naturally occurring amino acid residue, but in some cases non-naturally occurring amino acid residues may also be substituted. Non-natural amino acids, as the term is used herein, are non-proteinogenic (i.e., non-protein coding) amino acids that either occur naturally or are chemically synthesized. Examples include but are not limited to β-amino acids (β3 and β2), homo-amino acids, proline and pyruvic acid derivatives, 3-substituted alanine derivatives, glycine derivatives, ring-substituted phenylalanine and tyrosine derivatives, linear core amino acids, diarnino acids, D-amino acids, and N-methyl amino acids. In some embodiments, the amino acid can be substituted or unsubstituted. The substituted amino acid or substituent can be a halogenated aromatic or aliphatic amino acid, a halogenated aliphatic or aromatic modification on the hydrophobic side chain, or an aliphatic or aromatic modification.

The "percent identity" of two amino acid sequences is determined using the algorithm of Karlin and Altschul Proc. Natl. Acad. Sci. USA 87:2264-68, 1990, modified as in Karlin and Altschul Proc. Natl. Acad. Sci. USA 90:5873-77, 1993. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. J. Mol. Biol. 215:403-10, 1990. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to the protein molecules of interest. Where gaps exist between two sequences, Gapped BLAST can be utilized as described in Altschul et al, Nucleic Acids Res. 25(17):3389-3402, 1997. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used.

Accordingly, some aspects of the present disclosure provide isolated BoNT polypeptides. In some embodiments, the isolated BoNT polypeptide is a full-length BoNT/X
polypeptide. In some embodiments, the isolated BoNT polypeptide comprise the a amino acid sequence of SEQ ID NO: 1. In some embodiments, the isolated BoNT/X polypeptide comprises an amino acid sequence that has at least 85% identity to SEQ ID NO: 1. For example, the isolated BoNT polypeptide may comprise an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 1. In some embodiments, the isolated BoNT polypeptide consists of the amino acid sequence of SEQ ID NO: 1. In some embodiments, the isolated BoNT polypeptide is an X-LC-HN polypeptide. In some embodiments, the isolated BoNT polypeptide comprise the a amino acid sequence of SEQ ID NO: 2. In some embodiments, the isolated BoNT polypeptide comprises an amino acid sequence that has at least 85% identity to SEQ ID NO: 2. For example, the isolated BoNT polypeptide may comprise an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 2. In some embodiments, the isolated BoNT polypeptide comprises an amino acid sequence that has 85%, 86%, 87%, 88%, 89% ,90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identity to SEQ ID NO: 2. In some embodiments, the isolated BoNT polypeptide consists of the amino acid sequence of SEQ ID NO: 2.

In some embodiments, the isolated BoNT polypeptide is an X-LC polypeptide. In some embodiments, the isolated BoNT polypeptide comprise the a amino acid sequence of SEQ ID NO: 3. In some embodiments, the isolated BoNT polypeptide comprises an amino acid sequence that has at least 85% identity to SEQ ID NO: 3. For example, the isolated BoNT polypeptide may comprise an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 3. In some embodiments, the isolated BoNT polypeptide comprises an amino acid sequence that has 85%, 86%, 87%, 88%, 89% ,90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identity to SEQ ID NO: 3. In some embodiments, the isolated BoNT polypeptide consists of the amino acid sequence of SEQ ID NO: 3.
The X-LC polypeptide may be introduced alone into cells where the cleavage of a BoNT substrate (e.g., a SNARE protein) is desired for research or therapeutic purpose, by any known techniques of expression an exogenous protein in the art, e.g., transfection of LC coding sequence directly into cells, via lentiviral vectors, via AAV vectors, or fusing X-LC with cell penetrating peptides.

In some embodiments, the BoNT polypeptides of the present disclosure is a full-length BoNT/X comprising a protease domain (LC), a linker region, a translocation domain (H_N), and a receptor binding domain (He), wherein the linker region is located between the protease domain and the translocation domain. Like other BoNTs, BoNT/X is initially produced as a single polypeptide and is activated via the cleavage of the linker region between LC and H_N either bacterial or host proteases. This process is known as "activation" and is essential for the activity of BoNT/X. After the cleavage, the LC and H_N remain connected via an inter-chain disulfide bond prior to translocation of LC into the cytosol of cells, where the disulfide bond is reduced in order to release the LC into the cytosol. BoNT/X contains two cysteine that are conserved compared to other BoNTs, C423 and C467. Interestingly, BoNT/X also contains an additional cysteine (C461), which is unique to BoNT/X. The formation of the inter-chain disulfide bond (C423-C461, or C423-C467) is required for BoNT/X activity.

In addition to the cysteines in the linker region, the receptor binding domain of BoNT contains another cysteine, C1240, which can also form inter-molecular disulfide bonds with other cysteines in BoNT/X. These intermolecular disulfide bonds causes BoNT/X to aggregate and destabilizes the protein (FIG. 4B). Replacing the cysteines that are not required for BoNT/X activity may produces BoNT/X polypeptides with increased stability.

Accordingly, some aspects of the present disclosure provide modified BoNT/X polypeptide comprising one or more substitution mutation(s) in C461, C467, or C1240, which are more stable than the wild-type BoNT/X and have comparable activities. The cysteines may be substituted with any amino acids that abolish the formation of disulfide bonds. In some embodiments, the cysteines are substituted with serine (S) or alanine (A). Possible combinations of substitution mutations that may be present in the modified BoNTs of the present disclosure are, without limitation: C461S, C461A, C467S, C467A, C1240S, C1240A, C461S/C1240S, C461A/C1240S, C461S/C1240A, C467A/C1240A, C467S/C1240S, C467A/C1240S, C467S/C1240A, and C467A/C1240A. "/" indicates double mutations. In some embodiments, the modified BoNT/X polypeptide of the present disclosure comprises an amino acid sequence of any one of SEQ ID NOs: 4-17. In some embodiments, the modified
BoNT/X polypeptide comprises an amino acid sequence that has at least 85% identity to any one of SEQ ID NO: 4-17, and does not have the amino acid sequence of SEQ ID NO: 1. For example, the modified BoNT/X polypeptide may comprise an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any one of SEQ ID NOs: 4-17, and does not have the amino acid sequence of SEQ ID NO: 1. In some embodiments, the modified BoNT/X polypeptide comprises an amino acid sequence that has 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identity to any one of SEQ ID NOs: 4-17, and does not have the amino acid sequence of SEQ ID NO: 1. In some embodiments, the modified BoNT/X polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 4-17.

In some embodiments, the modified BoNT polypeptide of the present disclosure is a modified BoNT/X-LC-H_N polypeptide comprising the substitution mutations described herein. In some embodiments, the modified BoNT/X-LC-H_N comprises one single substitution mutation in a position corresponding to C461 or C467 in SEQ ID NO: 2. In some embodiments, the modified BoNT/X-LC-HN comprises one single substitution mutation corresponding to C461A, C461S, C467A, or C467S in SEQ ID NO: 2. In some embodiments, the modified BoNT/X polypeptide of the present disclosure comprises an amino acid sequence of any one of SEQ ID NOs: 18-21. In some embodiments, the modified BoNT/X-LC-H_N polypeptide comprises an amino acid sequence that has at least 85% identity to any one of SEQ ID NO: 18-21, and does not have the amino acid sequence of SEQ ID NO: 2. For example, the modified BoNT/X-LC-H_N polypeptide may comprise an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any one of SEQ ID NOs: 18-21, and does not have the amino acid sequence of SEQ ID NO: 2. In some embodiments, the modified BoNT/X-LC-H_N polypeptide comprises an amino acid sequence that has 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identity to any one of SEQ ID NOs: 18-21, and does not have the amino acid sequence of SEQ ID NO: 2. In some embodiments, the modified BoNT/X-LC-H_N polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 18-21.

The modified BoNT polypeptide comprising one or more substitution mutation(s) (e.g.,
in C461, C467, or C1240) described herein does not form inter-molecular disulfide bonds that cause aggregation of the protein, and are therefore more stable than their corresponding wild type proteins. The activity of the BoNT polypeptides are not affected by the substitution mutations in the cysteines. Thus, the modified BoNT/X may be more suitable for therapeutic use than the wild type BoNT/X due to its increased stability.

Other aspects of the present disclosure provide chimeric BoNTs comprising BoNT/X-LC-HN described herein and the receptor binding domain (He) from a different BoNT. For example, the receptor binding domain may be from any one of BoNT/A, BoNT/B, BoNT/C, BoNT/D, BoNT/E, BoNT/F, and BoNT/G. Thus, the chimeric BoNTs contemplated herein include BoNT/X-LC-H_N-A-H_c, BoNT/X-LC-H_N-B-H_c, BoNT/X-LC-H_N-C-H_c, BoNT/X-LC-HN-D-HC, BoNT/X-LC-HN-E-HC, BoNT/X-LC-H_N-F-H_c, and BoNT/X-LC-H_N-G-Hc. It is to be understood that the He domain of any subtypes of the seven known serotypes (e.g., A, B, C, D, E, F, or G) are suitable for the chimeric toxin. When BoNT/A, BoNT/B, BoNT/C, BoNT/D, BoNT/E, BoNT/F, or BoNT/G is referred to, it encompasses all the subtypes. For example, BoNT/A has 8 subtypes, BoNT/A1, BoNT/A2, BoNT/A3, BoNT/A4, BoNT/A5, BoNT/A6, BoNT/A7, or BoNT/A8, and the H_c of any one of these BoNT/A subtypes are suitable for use in the chimeric BoNT of the present disclosure. Similarly, the He of any one of the 8 subtypes of BoNT/B, i.e., BoNT/B1, BoNT/B2, BoNT/B3, BoNT/B4, BoNT/B5, BoNT/B6, BoNT/B7, or BoNT/B8, are suitable for use in the chimeric BoNT of the present disclosure.

In some embodiments, BoNT/X-LC-H_N-AI-H_c (SEQ ID NO: 22), BoNT/X-LC-H_N-BI-Hc (SEQ ID NO: 23), and BoNT/X-LC-H_N-Cl-H_c (SEQ ID NO: 24) are provided. In some embodiments, the chimeric BoNT polypeptide comprises an amino acid sequence that has at least 85% identity to any one of SEQ ID NO: 22-24. For example, the chimeric BoNT polypeptide may comprise an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any one of SEQ ID NOs: 22-24. In some embodiments, the chimeric BoNT polypeptide comprises an amino acid sequence that has 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identity to any one of SEQ ID NOs: 22-24. In some embodiments, the chimeric BoNT polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 22-24.
modified BoNT/X-LC-H \textsubscript{N} comprising a substitution mutation in the linker region, \textit{e.g.}, in a position corresponding to C461 or C467 of SEQ ID NO: 2. For example, the BoNT/X-LC-H \textsubscript{N} in the chimeric BoNT may comprise a substitution mutation corresponding to C461A, C467A, C461S, or C467S of SEQ ID NO: 2. For example, the chimeric BoNT polypeptide of the present disclosure may comprise an amino acid sequence of any one of SEQ ID NOs: 25-30. In some embodiments, the chimeric BoNT polypeptide comprises an amino acid sequence that has at least 85\% identity to any one of SEQ ID NO: 25-30. For example, the chimeric BoNT polypeptide may comprise an amino acid sequence that has at least 85\%, at least 86\%, at least 87\%, at least 88\%, at least 89\%, at least 90\%, at least 91\%, at least 92\%, at least 93\%, at least 94\%, at least 95\%, at least 96\%, at least 97\%, at least 98\%, at least 99\%, or at least 99.5\% identity to any one of SEQ ID NOs: 25-30. In some embodiments, the chimeric BoNT polypeptide comprises an amino acid sequence that has 85\%, 86\%, 87\%, 88\%, 89\%, 90\%, 91\%, 92\%, 93\%, 94\%, 95\%, 96\%, 97\%, 98\%, 99\%, 99.5\%, or 100\% identity to any one of SEQ ID NOs: 25-30. In some embodiments, the chimeric BoNT polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 25-30.

To generate the chimeric toxins, \textit{e.g.}, the BONT/X-LC-HN-A 1-HC toxin, the X-LC-HN fragment comprising amino acid of about 1-892 (SEQ ID NO: 2) is fused to the receptor binding domain of any one of BoNT/A, BoNT/B, BoNT/C, BoNT/D, BoNT/E, BoNT/F, and BoNT/G. The receptor binding domains of different BoNTs correspond to amino acids of about 860-1291 of BoNT/B 1. It is to be understood that the border of the X-LC-H\textsubscript{N} fragment and/or the receptor binding domains may vary by 1-25 amino acids. For example, the X-LC-H\textsubscript{N} fragment that may be used for the chimeric toxin may comprise amino acids 1-917 or 1-867 of BoNT/X. In some embodiments, the X-LC-H\textsubscript{N} fragment that may be used for the chimeric toxin may comprise amino acids 1-893, 1-894, 1-895, 1-896, 1-897, 1-898, 1-900, 1-901, 1-902, 1-892, 1-891, 1-890, 1-889, 1-888, 1-887, 1-886, 1-885, 1-884, or 1-883 of BoNT/X. Similarly, the receptor binding that may be used for the chimeric toxin may comprise amino acid corresponding to 885-1291 or 835-1291 of BoNT/X. In some embodiments, the receptor binding that may be used for the chimeric toxin may comprise amino acid corresponding to 860-1291, 861-1291, 862-1291, 863-1291, 864-1291, 865-1291, 866-1291, 867-1291, 868-1291, 869-1291, 870-1291, 860-1291, 859-1291, 858-1291, 857-1291, 856-1291, 855-1291, 854-1291, 853-1291, 852-1291, or 851-1291 of BoNT/B. The skilled artisan is able to identified the domains that may be used for the chimeric toxin of the present disclosure, based on his/her knowledge in protein homology, with or without the assistance of
a sequence alignment software. The methods of fusing the fragments are standard recombinant techniques that are well known to one skilled in the art.

Further contemplated herein are modified BoNT/X polypeptides comprising a modified linker region, wherein the linker region comprises a specific protease cleavage site. A "specific protease cleavage site," as used herein, refers to a recognition and cleavage site for a specification protease, as opposed to a sequence that is recognized and cleavage by more than one non-specific proteases. Such specific proteases include, without limitation: thrombin, TEV, PreScission, Factor Xa, MMP-12, MMP-13, MMP-17, MMP-20, Granzyme-B, and Enterokinase. The cleavage site of the specific proteases may be added to the linker region of the BoNT/X polypeptide via insertion or replacement of the existing amino acids in the linker region (e.g., replace amino acids 424-460 of the BoNT/X polypeptide). The sequences of the specific protease cleavage sites sequences are also provided: LVPRIGS (thrombin, SEQ ID NO: 50), ENLYFQIG (TEV, SEQ ID NO: 51), LEVLFQIGP (PreScission, SEQ ID NO: 52), IEGRI or IDGRI (Factor Xa, SEQ ID NO: 53 or 54), DDDDKI (Enterokinase, SEQ ID NO: 55) and AHREQIGGI (SUMO protease, SEQ ID NO: 56). "I" indicates where cleavage occurs.

Other aspects of the present disclosure provide the functional characterization of the BoNT/X polypeptides. The BoNT/X polypeptides, modified BoNT/X polypeptides, and chimeric BoNT polypeptides of the present disclosure can bind and enter target cells, e.g., neurons, and cleave its substrate proteins, e.g. SNARE proteins. The term "SNARE proteins," as used herein, refers to SNAP (Soluble NSF Attachment Protein) Receptors, which is a large protein superfamily consisting of more than 60 members in yeast and mammalian cells. The primary role of SNARE proteins is to mediate vesicle fusion, i.e., the fusion of vesicles with their target membrane bound compartments (such as a lysosome). The best studied SNARE proteins are those that mediate docking of synaptic vesicles with the presynaptic membrane in neurons, e.g., SNAP-25, VAMP1, VAMP2, VAMP3, VAMP4, VAMP5, VAMP7, VAMP8, syntaxinl, and Ykt6. Several of these SNARE proteins are substrates of BoNTs. For example, VAMP1, VAMP2, VAMP3, SNAP-25, and syntaxin 1 have been shown to be cleaved by known BoNTs, e.g., BoNT/A and BoNT/B.

Provided herein are data showing that BoNT/X cleaves the SNARE proteins that are known substrates of BoNTs. One surprising finding of the present disclosure is that BoNT/X is able to cleave several SNARE proteins that other BoNTs are not able to cleave, e.g., VAMP4, VAMP5, and Ykt6. VAMP4 is widely expressed and is known to mediate vesicle fusion between trans-Golgi network (TGN) and endosomes, as well as homotypic fusion of
endosomes. BoNTs are traditionally known to be limited to target SNAREs that mediate vesicle exocytosis onto plasma membranes. BoNT/X is the first BoNT that is capable of cleaving SNAREs mediating other type fusion events inside cells that is not with plasma membrane as the destine. VAMP4 may also contribute to asynchronous synaptic vesicle exocytosis, enlargeosome exocytosis, and activity-dependent bulk endocytosis (ADBE) in neurons. In addition, VAMP4 has been implicated in granule release in immune cells. Thus, BoNT/X might have a unique potential among all BoNTs to modulate inflammatory secretion in immune cells, which can be exploited therapeutically. VAMP5 is mainly expressed in muscles and its function remains to be established. BoNT/X will be a unique tool for investigating the function of VAMP4 and VAMP5. Ykt6 functions in endoplasmic reticulum to Golgi transport. It also functions in early/recycling endosome to TGN transport. The identification of Ykt6 as a substrate of the BoNT polypeptides described herein is significant because it opens up new therapeutic application for blocking secretion in a wide range of cells by BoNTs.

Another surprising finding of the present disclosure is that BoNT/X cleaves the SNARE proteins at a novel site what was not previously described. As illustrated in the Examples and Figures of the present disclosure, BoNT/X cleaves between amino acids R66-S67 in VAMPI, VAMP2, and VAMP3. R66-A67 is a novel cleavage site distinct from established target sites for all other BoNTs (FIG. 2F). It is also the only BoNT cleavage site located within a region previously known as the SNARE motif (FIG. 2F).

Accordingly, the BoNT polypeptides of the present disclosure have expanded profile of target cells and substrates. In some embodiments, the BoNT polypeptide cleaves a SNARE protein in the cell. In some embodiments, the BoNT polypeptide cleaves a SNARE protein selected from the group consisting of: SNAP-25, VAMPI, VAMP2, VAMP3, VAMP4, VAMP5, Ykt6, and syntaxin 1. In some embodiments, the BoNT polypeptide cleaves VAMPI (SEQ ID NO: 39). In some embodiments, the BoNT polypeptide cleaves VAMPI between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 39. In some embodiments, the BoNT polypeptide cleaves VAMP2 (SEQ ID NO: 40). In some embodiments, the BoNT polypeptide cleaves VAMP2 between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 40. In some embodiments, the BoNT polypeptide cleaves VAMP3 (SEQ ID NO: 31). In some embodiments, the BoNT polypeptide cleaves VAMP3 between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 41. In some embodiments, the BoNT polypeptide cleaves VAMP4 (SEQ ID NO: 42). In some embodiments, the BoNT polypeptide
cleaves VAMP4 between amino acid residues corresponding to K87 and S88 of SEQ ID NO: 42. In some embodiments, the BoNT polypeptide cleaves VAMP5 (SEQ ID NO: 43). In some embodiments, the BoNT polypeptide cleaves VAMP5 between amino acid residues corresponding to R40 and S41 of SEQ ID NO: 43. In some embodiments, the BoNT polypeptide cleaves Ykt6 (SEQ ID NO: 44). In some embodiments, the BoNT polypeptide cleaves Ykt6 between amino acid residues corresponding to K173 and S174 of SEQ ID NO: 44.

In some embodiments, the BoNT polypeptide of the present disclosure cleaves a SNARE protein in a target cell. As used herein, a "target cell" means a cell that is a naturally occurring cell that BoNT is capable of entering or intoxicating. In some embodiments, a target cell is a secretory cell, e.g., a neuron or a secretory immune cell. Examples of neurons that may be BoNT target cells include, without limitation, motor neurons; sensory neurons; autonomic neurons; such as, e.g., sympathetic neurons and parasympathetic neurons; non-peptidergic neurons, such as, e.g., cholinergic neurons, adrenergic neurons, noradrenergic neurons, serotonergic neurons, GABAergic neurons; and peptidergic neurons, such as, e.g., Substance P neurons, Calcitonin Gene Related Peptide neurons, vasoactive intestinal peptide neurons, Neuropeptide Y neurons, cholecystokinin neurons.

The BoNT polypeptide of the present disclosure, e.g., the BoNT/X or the modified BoNT/X polypeptide, is able to target other types of secretory cells other than neurons, due to its ability to cleave VAMP4 or Ykt6. In some embodiments, the secretory cell targeted by the BoNT polypeptide is a secretory immune cell. A "secretory immune cell," as used herein, refers to immune cells that secrets cytokines, chemokines, or antibodies. Such secretory immune cells may be innate immune cells including, without limitation, natural killer cells, mast cells, eosinophils, basophils, macrophages, neutrophils, and dendritic cells. Secretory immune cells that secret antibodies (e.g., white blood cells) may also be targeted by the BoNT polypeptides of the present disclosure. Non-limiting examples of antibody secreting cells include, without limitation, plasma B cells, plasmocytes, plasmacytes, and effector B cells. In some embodiments, the target cell is a cultured cell, e.g., a cultured neuron or a cultured secretory immune cell. In some embodiments, the target cell is in vivo. In some embodiments, target cell is from a mammal. In some embodiments, the mammal is a human. In embodiments, the mammal is a rodent, e.g., a mouse or a rat.

In some embodiments, the BoNT polypeptide suppresses neuronal activity. In some embodiments, the BoNT polypeptide modulates immune response. In some embodiments, the
BoNT polypeptide induces flaccid paralysis. "Flaccid paralysis" refers to a clinical manifestation characterized by weakness or paralysis and reduced muscle tone without other obvious cause (e.g., trauma).

Other aspects of the present disclosure provide modified BoNT/X polypeptides comprising an inactive protease domain. Such BoNT/X polypeptides (also referred to herein as "inactive BoNT/X") can enter the target cells but cannot cleave the substrate proteins (e.g., a SNARE protein) due to the inactivation of the protease domain. In some embodiments, the inactive BoNT/X is an **X-LC-HN** fragment comprising: a) an inactive protease domain; b) a linker region; and c) a translocation domain. In some embodiments, the inactive BoNT/X is a full length BoNT/X polypeptide comprising: a) an inactive protease domain; b) a linker region; c) a translocation domain; and d) a receptor binding domain. In some embodiments, the inactive protease domain comprises one or more substitution mutation(s) in a position corresponding to R360, Y363, H227, E228, or H231 of SEQ ID NO: 1. In some embodiments, the one or more substitution mutation(s) corresponds to R360A/Y363F, H227Y, E228Q, or H231Y in SEQ ID NO: 1. It is to be understood that the inactive BoNT/X polypeptide may comprise any mutation(s) that inactivates the protease domain.

In some embodiments, the inactive BoNT/X polypeptide comprises an amino acid sequence of any one of SEQ ID NOs: 31-38. In some embodiments, the inactive BoNT/X polypeptide comprises an amino acid sequence that has at least 85% identity to any one of SEQ ID NOs: 31-38. For example, the inactive BoNT/X polypeptide may comprise an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any one of SEQ ID NOs: 31-38. In some embodiments, the inactive BoNT/X polypeptide comprises an amino acid sequence that has 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, or 100% identity to any one of SEQ ID NOs: 31-38. In some embodiments, the inactive BoNT/X polypeptide consists of the amino acid sequence of any one of SEQ ID NOs: 31-38.

In some embodiments, the inactive BoNT/X (e.g., inactive **X-LC-HN** or inactive full length BoNT/X) further comprises mutations in the linker region. In some embodiments, the modification in the linker region comprises one single substitution mutation in a position corresponding to C461 or C467 of SEQ ID NO: 1. In some embodiments, the single substitution mutation corresponds to C461A, C461S, C467A, or C467S in SEQ ID NO: 1. In some embodiments, the inactive BoNT/X (e.g., the inactive full length BoNT/X) further
comprises a modification in the receptor binding domain. In some embodiments, the modification in the receptor binding domain comprises a substitution mutation in a position corresponding to C1240 of SEQ ID NO: 1.

It is also envisioned that the modified BoNT/X polypeptide comprising an inactive protease domain described herein can be utilized as a delivery tool to target cells (e.g., neurons) in humans. For example, the modified BoNT/X can be linked to other therapeutic agents, covalently or non-covalently, and acts as the targeting vehicle to deliver the therapeutic agents to target cells in humans.

As such, another aspect of the disclosure relates to a chimeric polypeptide molecule comprising a first portion that is an inactive BoNT/X, comprising one or more substitution mutations that inactivates the protease domain, linked to a second portion. The second portion of the molecule can be a bioactive molecule such as a therapeutic agent (e.g., a polypeptide or non-polypeptide drug). Linkage of the first and second portions of the molecule can be covalent (e.g., in the form of a fusion protein) or non-covalent. Methods of such linkage are known in the art and can readily be applied by the skilled practitioner. When the second portion of the chimeric molecule is a polypeptide and the chimeric molecule is in the form of a protein, nucleic acids and nucleic acid vectors encoding such chimeric molecules are provided.

Also provided are cells comprising the nucleic acids or nucleic acid vectors, and cells expressing such chimeric molecules. The chimeric molecules in a fusion protein form may be expressed and isolated using the methods disclosed herein.

The modified BoNT/X polypeptides, the chimeric BoNT polypeptides, or the chimeric molecules comprising a second portion that is a polypeptide of the present disclosure (e.g., without limitation, polypeptides comprising amino acid sequence of any one of SEQ ID NOs: 1-38), will generally be produced by expression form recombinant nucleic acids in appropriate cells (e.g., E. coli, or insect cells) and isolated. The nucleic acids encoding the polypeptides described herein may be obtained, and the nucleotide sequence of the nucleic acids determined, by any method known in the art.

Further provided herein are isolated and/or recombinant nucleic acids encoding any of the BoNT polypeptides disclosed herein. The nucleic acids encoding the isolated polypeptide fragments of the present disclosure, may be DNA or RNA, double-stranded or single stranded. In certain aspects, the subject nucleic acids encoding the isolated polypeptide fragments are further understood to include nucleic acids encoding polypeptides that are variants of any one of the modified BoNT polypeptides described herein.
Variant nucleotide sequences include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants. In some embodiments, the isolated nucleic acid molecule of the present disclosure comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity of any one of SEQ ID NOs: 1-38. In some embodiments, the isolated nucleic acid molecule of the present disclosure comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence that has 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity of any one of SEQ ID NOs: 1-38.

In some embodiments, the nucleic acid is comprised within a vector, such as an expression vector. In some embodiments, the vector comprises a promoter operably linked to the nucleic acid.

A variety of promoters can be used for expression of the polypeptides described herein, including, but not limited to, cytomegalovirus (CMV) intermediate early promoter, a viral LTR such as the Rous sarcoma virus LTR, HIV-LTR, HTLV-1 LTR, the simian virus 40 (SV40) early promoter, E. coli lac UV5 promoter, and the herpes simplex tk virus promoter. Regulatable promoters can also be used. Such regulatable promoters include those using the lac repressor from E. coli as a transcription modulator to regulate transcription from lac operator-bearing mammalian cell promoters [Brown, M. et al, Cell, 49:603-612 (1987)], those using the tetracycline repressor (tetR) [Gossen, M., and Bujard, H., Proc. Natl. Acad. Sci. USA 89:5547-5551 (1992); Yao, F. et al, Human Gene Therapy, 9:1939-1950 (1998); Shockelt, P., et al, Proc. Natl. Acad. Sci. USA, 92:6522-6526 (1995)].

Other systems include FK506 dimer, VP16 or p65 using estradiol, RU486, diphenol murslerone, or rapamycin. Inducible systems are available from Invitrogen, Clontech and Ariad. Regulatable promoters that include a repressor with the operon can be used. In one embodiment, the lac repressor from Escherichia coli can function as a transcriptional modulator to regulate transcription from lac operator-bearing mammalian cell promoters [M. Brown et al, Cell, 49:603-612 (1987)]; Gossen and Bujard (1992); [M. Gossen et al, Natl. Acad. Sci. USA, 89:5547-5551 (1992)] combined the tetracycline repressor (tetR) with the transcription activator (VP 16) to create a tetR-mammalian cell transcription activator fusion protein, tTa (tetR-VP 16), with the tetO-bearing minimal promoter derived from the human cytomegalovirus (HCMV) major immediate-early promoter to create a tetR-tet operator system.
to control gene expression in mammalian cells. In one embodiment, a tetracycline inducible
switch is used (Yao et al., Human Gene Therapy; Gossen et al., Natl. Acad. Sci. USA,

Additionally, the vector can contain, for example, some or all of the following: a
selectable marker gene, such as the neomycin gene for selection of stable or transient
transfectants in mammalian cells; enhancer/promoter sequences from the immediate early gene
of human CMV for high levels of transcription; transcription termination and RNA processing
signals from SV40 for mRNA stability; SV40 polyoma origins of replication and ColEl for
proper episomal replication; internal ribosome binding sites (IREs), versatile multiple
cloning sites; and T7 and SP6 RNA promoters for in vitro transcription of sense and antisense
RNA. Suitable vectors and methods for producing vectors containing transgenes are well
known and available in the art.

An expression vector comprising the nucleic acid can be transferred to a host cell by
conventional techniques (e.g., electroporation, liposomal transfection, and calcium phosphate
precipitation) and the transfected cells are then cultured by conventional techniques to produce
the polypeptides described herein. In some embodiments, the expression of the polypeptides
described herein is regulated by a constitutive, an inducible or a tissue-specific promoter.

The host cells used to express the isolated polypeptides described herein may be either
bacterial cells such as Escherichia coli, or, preferably, eukaryotic cells. In particular,
mammalian cells, such as Chinese hamster ovary cells (CHO), in conjunction with a vector
such as the major intermediate early gene promoter element from human cytomegalovirus is an
effective expression system for immunoglobulins (Foecking et al. (1986) "Powerful And
Versatile Enhancer-Promoter Unit For Mammalian Expression Vectors," Gene 45:101-106;
Cockett et al. (1990) "High Level Expression Of Tissue Inhibitor Of Metalloproteinases In
Chinese Hamster Ovary Cells Using Glutamine Synthetase Gene Amplification,"
Biotechnology 8:662-667). A variety of host-expression vector systems may be utilized to
express the isolated polypeptides described herein. Such host-expression systems represent
vehicles by which the coding sequences of the isolated polypeptides described herein may be
produced and subsequently purified, but also represent cells which may, when transformed or
transfected with the appropriate nucleotide coding sequences, express the isolated polypeptides
described herein in situ. These include, but are not limited to, microorganisms such as bacteria
(e.g., E. coli and B. subtilis) transformed with recombinant bacteriophage DNA, plasmid DNA
or cosmid DNA expression vectors containing coding sequences for the isolated polypeptides
described herein; yeast (e.g., Saccharomyces pichia) transformed with recombinant yeast expression vectors containing sequences encoding the isolated polypeptides described herein; insect cell systems infected with recombinant virus expression vectors (e.g., baclovirus) containing the sequences encoding the isolated polypeptides described herein; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus (CaMV) and tobacco mosaic virus (TMV)) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing sequences encoding the isolated polypeptides described herein; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 293T, 3T3 cells, lymphotic cells (see U.S. Pat. No. 5,807,715), Per C.6 cells (human retinal cells developed by Crucell) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the polypeptides being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of polypeptides described herein, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the E. coli expression vector pUR278 (Riether et al. 1983) "Easy Identification Of cDNA Clones," EMBO J. 2:1791-1794), in which the coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye et al. 1985) "Up-Promoter Mutations In The lpp Gene Of Escherichia Coli," Nucleic Acids Res. 13:3101-3110; Van Heeke et al. (1989) "Expression Of Human Asparagine Synthetase In Escherichia Coli," J. Biol. Chem. 24:5503-5509); and the like, pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption and binding to a matrix glutathione-agarose beads followed by elution in the presence of free glutathione.

The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety. In an insect system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in Spodoptera frugiperda cells. The coding sequence may be cloned individually into non-essential regions (e.g., the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (e.g., the polyhedrin promoter).
In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the immunoglobulin molecule in infected hosts (e.g., see Logan et al. (1984) "Adenovirus Tripartite Leader Sequence Enhances Translation Of mRNAs Late After Infection," Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted antibody coding sequences. These signals include the ATG initiation codon and adjacent sequences. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic.

The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bitter et al. (1987) "Expression And Secretion Vectors For Yeast," Methods in Enzymol. 153:516-544). In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. For example, in certain embodiments, the polypeptides described herein may be expressed as a single gene product (e.g., as a single polypeptide chain, i.e., as a polyprotein precursor), requiring proteolytic cleavage by native or recombinant cellular mechanisms to form separate polypeptides described herein.

The disclosure thus encompasses engineering a nucleic acid sequence to encode a polypeptide precursor molecule comprising the polypeptides described herein, which includes coding sequences capable of directing post translational cleavage of said polypeptide precursor. Post-translational cleavage of the polypeptide precursor results in the polypeptides described herein. The post translational cleavage of the precursor molecule comprising the polypeptides described herein may occur in vivo (i.e., within the host cell by native or recombinant cell systems/mechanisms, e.g. furin cleavage at an appropriate site) or may occur in vitro (e.g., incubation of said polypeptide chain in a composition comprising proteases or peptidases of
known activity and/or in a composition comprising conditions or reagents known to foster the desired proteolytic action).

Purification and modification of recombinant proteins is well known in the art such that the design of the polyprotein precursor could include a number of embodiments readily appreciated by a skilled worker. Any known proteases or peptidases known in the art can be used for the described modification of the precursor molecule, e.g., thrombin or factor Xa (Nagai et al. (1985) "Oxygen Binding Properties Of Human Mutant Hemoglobins Synthesized In Escherichia Coli," Proc. Nat. Acad. Sci. USA 82:7252-7255, and reviewed in Jenny et al. (2003) "A Critical Review Of The Methods For Cleavage Of Fusion Proteins With Thrombin And Factor Xa," Protein Expr. Purif. 31:1-11, each of which is incorporated by reference herein in its entirety)), enterokinase (Collins-Racie et al. (1995) "Production Of Recombinant Bovine Enterokinase Catalytic Subunit In Escherichia Coli Using The Novel Secretory Fusion Partner DsbA," BiotechHology 13:982-987 hereby incorporated by reference herein in its entirety)), furin, and AcTEV (Parks et al. (1994) "Release Of Proteins And Peptides From Fusion Proteins Using A Recombinant Plant Virus Proteinase," Anal. Biochem. 216:413-417 hereby incorporated by reference herein in its entirety)) and the Foot and Mouth Disease Virus Protease C3.

Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERY, BHK, HeLa, COS, MDCK, 293, 293T, 3T3, WI38, BT483, Hs578T, HTB2, BT20 and T47D, CRL7030 and Hs578Bst.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express polypeptides described herein may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and
allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the polypeptides described herein. Such engineered cell lines may be particularly useful in screening and evaluation of polypeptides that interact directly or indirectly with the polypeptides described herein.


Methods commonly known in the art of recombinant DNA tecH₉iqlology which can be used are described in Ausubel et al. (eds.), 1993, Current Protocols in Molecular Biology, JoH₉ Wiley & Sons, NY; Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY; and in Chapters 12 and 13, Dracopoli et al. (eds), 1994, Current Protocols in

The expression levels of polypeptides described herein can be increased by vector amplification (for a review, see Bebbington and Hentschel, The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning, Vol. 3 (Academic Press, New York, 1987). When a marker in the vector system expressing a polypeptide described herein is amplifiable, increase in the level of inhibitor present in culture of host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the nucleotide sequence of a polypeptide described herein or a polypeptide described herein, production of the polypeptide will also increase (Crouse et al. (1983) "Expression And Amplification Of Engineered Mouse Dihydrofolate Reductase Minigenes," Mol. Cell. Biol. 3:257-266).

Once a polypeptide described herein has been recombinantly expressed, it may be purified by any method known in the art for purification of polypeptides, polypeptides or antibodies \(\text{e.g.},\) analogous to antibody purification schemes based on antigen selectivity) for example, by chromatography \(\text{e.g.},\) ion exchange, affinity, particularly by affinity for the specific antigen (optionally after Protein A selection where the polypeptide comprises an Fc domain (or portion thereof)), and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of polypeptides or antibodies.

Other aspects of the present disclosure relate to a cell comprising a nucleic acid described herein or a vector described herein.

The cell may be a prokaryotic or eukaryotic cell. In some embodiments, the cell in a mammalian cell. Exemplary cell types are described herein. Other aspects of the present disclosure related to a cell expressing the modified BoNT polypeptides described herein. The cell may be a prokaryotic or eukaryotic cell. In some embodiments, the cell in a mammalian cell. Exemplary cell types are described herein. The cell can be for propagation of the nucleic acid or for expression of the nucleic acid, or both. Such cells include, without limitation, prokaryotic cells including, without limitation, strains of aerobic, microaerophilic, capnophilic, facultative, anaerobic, gram-negative and gram-positive bacterial cells such as those derived from, \(\text{e.g.},\) Escherichia coli, Bacillus subtilis, Bacillus licheniformis, Bacteroides fragilis, Clostridia perfringens, Clostridia difficile, Caulobacter crescentus, Lactococcus lactis, Methylobacterium extorquens, Neisseria meningitidis, Neisseria meningitidis, Pseudomonas fluorescens and Salmonella typhimurium; and eukaryotic cells including, without limitation,
yeast strains, such as, e.g., those derived from Pichia pastoris, Pichia methanolica, Pichia angusta, Schizosaccharomyces pombe, Saccharomyces cerevisiae and Yarrowia lipolytica; insect cells and cell lines derived from insects, such as, e.g., those derived from Spodoptera frugiperda, Trichoplusia ni, Drosophila melanogaster and Manduca sexta; and mammalian cells and cell lines derived from mammalian cells, such as, e.g., those derived from mouse, rat, hamster, porcine, bovine, equine, primate and human. Cell lines may be obtained from the American Type Culture Collection, European Collection of Cell Cultures and the German Collection of Microorganisms and Cell Cultures. Non-limiting examples of specific protocols for selecting, making and using an appropriate cell line are described in e.g., INSECT CELL CULTURE ENGINEERING (Mattheus F. A. Goosen et al. eds., Marcel Dekker, 1993); INSECT CELL CULTURES: FUNDAMENTAL AND APPLIED ASPECTS (J. M. Vlak et al. eds., Kluwer Academic Publishers, 1996); Maureen A. Harrison & Ian F. Rae, GENERAL TECHNIQUES OF CELL CULTURE (Cambridge University Press, 1997); CELL AND TISSUE CULTURE: LABORATORY PROCEDURES (Alan Doyle et al eds., JoH_N Wiley and Sons, 1998); R. Ian FresH_Ney, CULTURE OF ANIMAL CELLS: A MANUAL OF BASIC TECHNIQUES (Wiley-Liss, 4.sup.th ed. 2000); ANIMAL CELL CULTURE: A PRACTICAL APPROACH (JoH_N R. W. Masters ed., Oxford University Press, 3.sup.rd ed. 2000); MOLECULAR CLONING A LABORATORY MANUAL, supra, (2001); BASIC CELL CULTURE: A PRACTICAL APPROACH (JoH_N M. Davis, Oxford Press, 2.sup.nd ed. 2002); and CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, supra, (2004).

These protocols are routine procedures within the scope of one skilled in the art and from the teaching herein. Yet other aspects of the present disclosure relate to a method of producing a polypeptide described herein, the method comprising obtaining a cell described herein and expressing nucleic acid described herein in said cell. In some embodiments, the method further comprises isolating and purifying a polypeptide described herein.

In some embodiments, botulinum neurotoxin can be obtained by establishing and growing cultures of Clostridium botulinum in a fermenter and then harvesting and purifying the fermented mixture in accordance with known procedures. All the botulinum toxin serotypes are initially synthesized as inactive single chain proteins which must be cleaved or nicked by proteases to become neuroactive.

The bacterial strains that make botulinum toxin serotypes A and G possess endogenous proteases and serotypes A and G can therefore be recovered from bacterial cultures in predominantly their active form. In contrast, botulinum toxin serotypes C, D and E are
synthesized by non-proteolytic strains and are therefore typically inactive when recovered from culture. Serotypes B and F are produced by both proteolytic and non-proteolytic strains and therefore can be recovered in either the active or inactive form. The proteolytic strains that produce, for example, the botulinum toxin type B serotype may only cleave a portion of the toxin produced. The production of BoNT/X polypeptides using these strains are contemplated herein.

The exact proportion of nicked to un-nicked molecules depends on the length of incubation and the temperature of the culture. Therefore, a certain percentage of a preparation of, for example, the botulinum toxin type B toxin may be inactive. In one embodiment, the neurotoxin of the present disclosure is in an active state. In one embodiment, the neurotoxin is in an inactive state. In one embodiment, a combination of active and inactive neurotoxin is envisioned.

One aspect of the present disclosure provides novel methods of producing BoNTs via an in vitro transpeptidase reaction that ligates two non-toxic fragments of BoNTs. Such methods comprise the steps of: (i) obtaining a first BoNT fragment comprising a light chain (LC) and a N-terminal domain of a heavy chain (HN), wherein the first BoNT fragment comprises a C-terminal LPXTGG (SEQ ID NO: 60) motif; (ii) obtaining a second BoNT fragment comprising a C-terminal domain of the heavy chain (HC); wherein the second BoNT fragment comprise a specific protease cleavage site at its N-terminus; (iii) cleaving the second BoNT fragment with a specific protease, wherein the cleavage results in a free glycine residue at the N-terminus; and (iv) contacting the first BoNT fragment and the second BoNT fragment in the presence of a transpeptidase, thereby ligating the first BoNT fragment and the second BoNT fragment to form a ligated BoNT.

In some embodiments, the first BoNT fragment comprises the X-LC-HN polypeptide described herein fused to a C-terminal LPXTGG (SEQ ID NO: 60) motif (e.g., SEQ ID NO: 45), or any variants thereof. In some embodiments, the second BoNT fragment comprises the He polypeptide described herein, or any variants thereof (e.g., SEQ ID NO: 46). It is understood that any BoNT fragments or domains may be ligated using the methods described herein.

The methods described herein may also be used to generate chimeric BoNTs. For example, the first BoNT fragment may be from BoNT serotype A, B, C, D, E, F, G, or X. Similarly, the second BoNT fragment may be from BoNT serotype A, B, C, D, E, F, G, or X. One skilled in the art will be able to discern the combinations that may be made. In some
embodiments, the chimeric BoNT polypeptides described herein (e.g., BoNT/X-LC-H \(_{N}\)-Al-H\(_{C}\), BoNT/X-LC-H \(_{N}\)-Bl-Hc, or BoNT/X-LC-H \(_{N}\)-Cl-H\(_{C}\)) are made using this method.

In some embodiments, the transpeptidase is a sortase. In some embodiments, the sortase is from Staphylococcus aureus (SrtA).

Other peptide ligation systems available in the art may also be used to ligate two non-toxic BoNT fragments. For example, an intein-mediated protein ligation reaction allows the ligation of a synthetic peptide or a protein with an N-terminal cysteine residue to the C-terminus of a bacterially expressed protein through a native peptide bond (Evans et al., 1998) Protein Sci.7, 2256-2264, Dawson et al.,(1994)Science266, 776-779; Tarn et al. (1995) Proc. Natl. Acad. Sci. USA92, 12485-12489, Muir et al.,(1998)Proc. Natl. Acad. Sci. USA95,6705-6710; Severinov and Muir(1998)J. Biol. Chem.273, 16205-16209, the entire contents of which are incorporated herein by references). Kits are commercially available (e.g., from New England Biolabs) for intern-mediated protein ligation reactions.

In some embodiments, the first BoNT fragment further comprises an affinity tag. In some embodiments, the affinity tag is fused to first BoNT fragment at the N-terminus. In some embodiments, the affinity tag is fused to the first BoNT fragment at the C-terminus. In the event that the affinity tag is fused to the C-terminus of the first BoNT fragment, the transpeptidase cleaves between the T and G in the LPXTGG (SEQ ID NO: 60) motif and removes the affinity tag before ligating the first BoNT fragment and the second BoNT fragment.

In some embodiments, the second BoNT fragment further comprises an affinity tag. In some embodiments, the affinity tag is fused to the first BoNT fragment at the N-terminus. In some embodiments, the affinity tag is fused to the second BoNT fragment at the C-terminus. In the event that the affinity tag is fused to the N-terminus of the first BoNT fragment, the specific protease cleaves in the specific protease cleavage site and removes the affinity tag before ligating the first BoNT fragment and the second BoNT fragment by the transpeptidase.

An "affinity tag," as used herein, refers to a polypeptide sequence that can bind specifically to a substance or a moiety, e.g., a tag comprising six Histidines bind specifically to Ni\(^{2+}\). Affinity tags may be appended to proteins to facilitate their isolation. The affinity tags are typically fused to proteins via recombinant DNA techniques known by those skilled in the art. The use of affinity tags to facilitate protein isolate is also well known in the art. Suitable affinity tags that may be used in accordance with the present disclosure include, without limitation, His6, GST, Avi, Strep, S, MBP, Sumo, FLAG, HA, Myc, SBP, E, Calmodulin,
Softag 1, Softag 3, TC, V5, VSV, Xpress, Halo, and Fc.

The second BoNT fragment has a specific protease cleavage at the N-terminus. Cleavage of the site by the specific protease results to a free glycine residue at the N-terminus of the second BoNT fragment. Suitable specific protease that may be used in accordance with the present disclosure include, without limitation: thrombin, TEV, PreScission, Enterokinase, and SUMO protease. In some embodiments, the specific protease is thrombin, and the cleavage site is :LVPRIGS (SEQ ID NO: 50).

The BoNT/X polypeptides described herein affords potential for therapeutic use. For example, BoNT/X might be more potent compared to other BoNT serotypes. BoNT/X is more versatile and may be more effective in a wide range of cells due to its ability to cleave more substrates than other BoNT serotypes.

Thus, the present disclosure also contemplates pharmaceutically compositions comprising the BoNT/X polypeptides or the chimeric molecules of the present disclosure. As it may also become clear later in the present disclosure, the pharmaceutical composition of the present disclosure, may further comprise other therapeutic agents suitable for the specific disease such composition is designed to treat. In some embodiments, the pharmaceutically composition of the present disclosure further comprises pharmaceutically-acceptable carriers.

The term "pharmaceutically-acceptable carrier", as used herein, means a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, manufacturing aid (e.g., lubricant, talc magnesium, calcium or zinc stearate, or steric acid), or solvent encapsulating material, involved in carrying or transporting the polypeptide from one site (e.g., the delivery site) of the body, to another site (e.g., organ, tissue or portion of the body).

A pharmaceutically acceptable carrier is "acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the tissue of the subject (e.g., physiologically compatible, sterile, physiologic pH, etc.). Some examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethylcellulose, methylcellulose, ethyl cellulose, microcrystalline cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) lubricating agents, such as magnesium stearate, sodium lauryl sulfate and talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as
propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol (PEG); (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) pH buffered solutions; (21) polyesters, polycarbonates and/or polyanhydrides; (22) bulking agents, such as polypeptides and amino acids (23) serum component, such as serum albumin, HDL and LDL; (22) C2-C12 alcohols, such as ethanol; and (23) other non-toxic compatible substances employed in pharmaceutical formulations. Wetting agents, coloring agents, release agents, coating agents, sweetening agents, flavoring agents, perfuming agents, preservative and antioxidants can also be present in the formulation. The terms such as "excipient", "carrier", "pharmaceutically acceptable carrier" or the like are used interchangeably herein. In some embodiments, a BoNT polypeptide of the present disclosure in a composition is administered by injection, by means of a catheter, by means of a suppository, or by means of an implant, the implant being of a porous, non-porous, or gelatinous material, including a membrane, such as a silastic membrane, or a fiber.

Typically, when administering the composition, materials to which the polypeptide of the disclosure does not absorb are used. In other embodiments, the BoNT polypeptides of the present disclosure are delivered in a controlled release system. Such compositions and methods for administration are provides in U.S. Patent publication No. 2007/0020295, the contents of which are herein incorporated by reference. In one embodiment, a pump may be used (see, e.g., Langer, 1990, Science 249:1527-1533; Sefton, 1989, CRC Crit. Ref. Biomed. Eng. 14:201; Buchwald et al., 1980, Surgery 88:507; Saudek et al., 1989, N. Engl. J. Med. 321:574).


The BoNT polypeptides of the present disclosure can be administered as pharmaceutical compositions comprising a therapeutically effective amount of a binding agent and one or more pharmaceutically compatible ingredients. In typical embodiments, the pharmaceutical composition is formulated in accordance with routine procedures as a
pharmaceutical composition adapted for intravenous or subcutaneous administration to a subject, e.g., a human being.

Typically, compositions for administration by injection are solutions in sterile isotonic aqueous buffer. Where necessary, the pharmaceutical can also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the pharmaceutical is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the pharmaceutical is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients can be mixed prior to administration. A pharmaceutical composition for systemic administration may be a liquid, e.g., sterile saline, lactated Ringer’s or Hank’s solution. In addition, the pharmaceutical composition can be in solid forms and re-dissolved or suspended immediately prior to use. Lyophilized forms are also contemplated. The pharmaceutical composition can be contained within a lipid particle or vesicle, such as a liposome or microcrystal, which is also suitable for parenteral administration. The particles can be of any suitable structure, such as unilamellar or plurilamellar, so long as compositions are contained therein.

The polypeptides of the present disclosure can be entrapped in ‘stabilized plasmid-lipid particles’ (SPLP) containing the fusogenic lipid dioleoylphosphatidylethanolamine (DOPE), low levels (5-10 mol %) of cationic lipid, and stabilized by a polyethyleneglycol (PEG) coating (Zhang Y. P. et al., Gene Ther. 1999, 6:1438-47). Positively charged lipids such as N-[1-(2,3-dioleoyloxi)propyl]-N,N,N-trimethyl-amoniummethylsulfate, or "DOTAP," are particularly preferred for such particles and vesicles. The preparation of such lipid particles is well known. See, e.g., U.S. Patent Nos. 4,880,635; 4,906,477; 4,911,928; 4,917,951; 4,920,016; and 4,921,757. The pharmaceutical compositions of the present disclosure may be administered or packaged as a unit dose, for example.

The term "unit dose" when used in reference to a pharmaceutical composition of the present disclosure refers to physically discrete units suitable as unitary dosage for the subject, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent; i.e., carrier, or vehicle. In some embodiments, the BoNT/X polypeptides described herein may be conjugated to a therapeutic moiety, e.g., an antibiotic. TecH₉iques for conjugating such therapeutic moieties
to polypeptides, including e.g., Fc domains, are well known; see, e.g., Amon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), 1985, pp. 243-56, Alan R. Liss, Inc.; Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), 1987, pp. 623-53, Marcel Dekker, Inc.; Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al. (eds.), 1985, pp. 475-506); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), 1985, pp. 303-16, Academic Press; and Thorpe et al. (1982) "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates," Immunol. Rev., 62:119-158. Further, the pharmaceutical composition can be provided as a pharmaceutical kit comprising (a) a container containing a polypeptide of the disclosure in lyophilized form and (b) a second container containing a pharmaceutically acceptable diluent (e.g., sterile water) for injection. The pharmaceutically acceptable diluent can be used for reconstitution or dilution of the lyophilized polypeptide of the disclosure. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In another aspect, an article of manufacture containing materials useful for the treatment of the diseases described above is included. In some embodiments, the article of manufacture comprises a container and a label.

Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers may be formed from a variety of materials such as glass or plastic. In some embodiments, the container holds a composition that is effective for treating a disease described herein and may have a sterile access port. For example, the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle. The active agent in the composition is an isolated polypeptide of the disclosure. In some embodiments, the label on or associated with the container indicates that the composition is used for treating the disease of choice. The article of manufacture may further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution, or dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, syringes, and package inserts with instructions for use.
The BoNT polypeptides (e.g., BoNT/X polypeptides), the chimeric molecules, and the pharmaceutical compositions of the present disclosure may be used for the treatment of conditions associated with unwanted neuronal activities. Thus, further provided herein are methods of treating a condition associated with unwanted neuronal activity, the method comprising administering a therapeutically effective amount of the BoNT polypeptide, the chimeric molecule, or the pharmaceutical composition described herein to thereby treat the condition. In some embodiments, the BoNT polypeptides, the chimeric molecules, and the pharmacetic compositions of the present disclosure contact one or more neuron(s) exhibiting unwanted neuronal activity.

Conditions typically treated with a neurotoxin (e.g., skeletal muscle conditions, smooth muscle conditions, glandular conditions, a neuromuscular disorder, an autonomic disorder, pain, or an aesthetic/cosmetic condition) are associated with unwanted neuronal activity, as determined by the skilled practitioner. Administration is by a route that contacts an effective amount of the composition to neurons exhibiting the unwanted activity. In some embodiments, the condition may be associated with overactive neurons or glands. Specific conditions envisioned for treatment by the methods discussed herein include, without limitation, spasmodic dysphonia, spasmodic torticollis, laryngeal dystonia, oromandibular dysphonia, lingual dystonia, cervical dystonia, focal hand dystonia, blepharospasm, strabismus, hemifacial spasm, eyelid disorder, cerebral palsy, focal spasticity and other voice disorders, spasmodic colitis, neurogenic bladder, anismus, limb spasticity, tics, tremors, bruxism, anal fissure, achalasia, dysphagia and other muscle tone disorders and other disorders characterized by involuntary movements of muscle groups, lacrimation, hyperhydrosis, excessive salivation, excessive gastrointestinal secretions as well as other secretory disorders, pain from muscle spasms, headache pain. In addition, the present disclosure can be used to treat dermatological or aesthetic/cosmetic conditions, for example, reduction of brow furrows, reduction of skin wrinkles.

One unique property of the BoNT/X polypeptides of the present disclosure is its ability to cleave VAMP4, VAMP5, and Ykt6. Thus, further contemplated herein are therapeutic use of the BoNT/X polypeptides in conditions associated with unwanted secretion activities in a wide range of cells. In some embodiments, the unwanted secretion is immune secretion. Conditions associated with unwanted immune secretion include, without limitation: inflammation, psoriasis, allergy, haemophagocytic lymphohistiocytosis, and alcoholic pancreatic disease.
The present disclosure can also be used in the treatment of sports injuries. Borodic U.S. Pat. No. 5,053,005 discloses methods for treating juvenile spinal curvature, *i.e.* scoliosis, using botulinum type A. The disclosure of Borodic is incorporated in its entirety herein by reference. In one embodiment, using substantially similar methods as disclosed by Borodic, a BoNT polypeptide can be administered to a mammal, preferably a human, to treat spinal curvature. In a suitable embodiment, a BoNT polypeptide comprising botulinum type E fused with a leucine-based motif is administered. Even more preferably, a BoNT polypeptide comprising botulinum type A-E with a leucine-based motif fused to the carboxyl terminal of its light chain is administered to the mammal, preferably a human, to treat spinal curvature.

In addition, the BoNT polypeptides can be administered to treat neuromuscular disorders using well known techniques that are commonly performed with botulinum type A. For example, the present disclosure can be used to treat pain, for example, headache pain, pain from muscle spasms and various forms of inflammatory pain. For example, Aoki U.S. Pat. No. 5,721,215 and Aoki U.S. Pat. No. 6,113,915 disclose methods of using botulinum toxin type A for treating pain. The disclosure of these two patents is incorporated in its entirety herein by reference.

Autonomic nervous system disorders can also be treated with a modified neurotoxin. For example, glandular malfunctioning is an autonomic nervous system disorder. Glandular malfunctioning includes excessive sweating and excessive salivation. Respiratory malfunctioning is another example of an autonomic nervous system disorder. Respiratory malfunctioning includes chronic obstructive pulmonary disease and asthma. Sanders *et al.* disclose methods for treating the autonomic nervous system; for example, treating autonomic nervous system disorders such as excessive sweating, excessive salivation, asthma, etc., using naturally existing botulinum toxins. The disclosure of Sander *et al.* is incorporated in its entirety by reference herein.

In one embodiment, substantially similar methods to that of Sanders *et al.* can be employed, but using a BoNT polypeptide, to treat autonomic nervous system disorders such as the ones discussed above. For example, a BoNT polypeptide can be locally applied to the nasal cavity of the mammal in an amount sufficient to degenerate cholinergic neurons of the autonomic nervous system that control the mucous secretion in the nasal cavity. Pain that can be treated by a modified neurotoxin includes pain caused by muscle tension, or spasm, or pain that is not associated with muscle spasm. For example, Binder in U.S. Pat. No. 5,714,468 discloses that headache caused by vascular disturbances, muscular tension, neuralgia and
neuropathy can be treated with a naturally occurring botulinum toxin, for example Botulinum type A. The disclosures of Binder are incorporated in its entirety herein by reference.

In one embodiment, substantially similar methods to that of Binder can be employed, but using a BoNT polypeptide described herein, to treat headache, especially the ones caused by vascular disturbances, muscular tension, neuralgia and neuropathy. Pain caused by muscle spasm can also be treated by an administration of a BoNT polypeptide described herein. For example, a botulinum type E fused with a leucine-based motif, preferably at the carboxyl terminal of the botulinum type E light chain, can be administered intramuscularly at the pain/spasm location to alleviate pain. Furthermore, a modified neurotoxin can be administered to a mammal to treat pain that is not associated with a muscular disorder, such as spasm.

In one broad embodiment, methods of the present disclosure to treat non-spasm related pain include central administration or peripheral administration of the BoNT polypeptide. For example, Foster et al. in U.S. Pat. No. 5,989,545 discloses that a botulinum toxin conjugated with a targeting moiety can be administered centrally (intrathecally) to alleviate pain. The disclosures of Foster et al. are incorporated in its entirety by reference herein.

In one embodiment, substantially similar methods to that of Foster et al. can be employed, but using the compositions described herein to treat pain. The pain to be treated can be an acute pain or chronic pain. An acute or chronic pain that is not associated with a muscle spasm can also be alleviated with a local, peripheral administration of the modified neurotoxin to an actual or a perceived pain location on the mammal.

In one embodiment, the BoNT polypeptide is administered subcutaneously at or near the location of pain, for example, at or near a cut. In some embodiments, the modified neurotoxin is administered intramuscularly at or near the location of pain, for example, at or near a bruise location on the mammal. In some embodiments, the BoNT polypeptide is injected directly into a joint of a mammal, for treating or alleviating pain caused by arthritic conditions. Also, frequent repeated injection or infusion of the modified neurotoxin to a peripheral pain location is within the scope of the present disclosure. Routes of administration for such methods are known in the art and easily adapted to the methods described herein by the skilled practitioner (e.g., see for example, Harrison's Principles of Internal Medicine (1998), edited by Anthony Fauci et al., 14.sup.th edition, published by McGraw Hill).

By way of non-limiting example, the treatment of a neuromuscular disorder can comprise a step of locally administering an effective amount of the molecule to a muscle or a group of muscles, the treatment of an autonomic disorder can comprise a step of locally
administering an effective of the molecule to a gland or glands, and the treatment of pain can comprise a step of administering an effective amount of the molecule the site of the pain. In addition, the treatment of pain can comprise a step of administering an effective amount of a modified neurotoxin to the spinal cord.

"A therapeutically effective amount" as used herein refers to the amount of each therapeutic agent of the present disclosure required to confer therapeutic effect on the subject, either alone or in combination with one or more other therapeutic agents. Effective amounts vary, as recognized by those skilled in the art, depending on the particular condition being treated, the severity of the condition, the individual subject parameters including age, physical condition, size, gender and weight, the duration of the treatment, the nature of concurrent therapy (if any), the specific route of administration and like factors within the knowledge and expertise of the health practitioner. These factors are well known to those of ordinary skill in the art and can be addressed with no more than routine experimentation. It is generally preferred that a maximum dose of the individual components or combinations thereof be used, that is, the highest safe dose according to sound medical judgment. It will be understood by those of ordinary skill in the art, however, that a subject may insist upon a lower dose or tolerable dose for medical reasons, psychological reasons or for virtually any other reasons. Empirical considerations, such as the half-life, generally will contribute to the determination of the dosage. For example, therapeutic agents that are compatible with the human immune system, such as polypeptides comprising regions from humanized antibodies or fully human antibodies, may be used to prolong half-life of the polypeptide and to prevent the polypeptide being attacked by the host's immune system.

Frequency of administration may be determined and adjusted over the course of therapy, and is generally, but not necessarily, based on treatment and/or suppression and/or amelioration and/or delay of a disease. Alternatively, sustained continuous release formulations of a polypeptide may be appropriate. Various formulations and devices for achieving sustained release are known in the art. In some embodiments, dosage is daily, every other day, every three days, every four days, every five days, or every six days. In some embodiments, dosing frequency is once every week, every 2 weeks, every 4 weeks, every 5 weeks, every 6 weeks, every 7 weeks, every 8 weeks, every 9 weeks, or every 10 weeks; or once every month, every 2 months, or every 3 months, or longer. The progress of this therapy is easily monitored by conventional techniques and assays.
The dosing regimen (including the polypeptide used) can vary over time. In some embodiments, for an adult subject of normal weight, doses ranging from about 0.01 to 1000 mg/kg may be administered. In some embodiments, the dose is between 1 to 200 mg. The particular dosage regimen, i.e., dose, timing and repetition, will depend on the particular subject and that subject's medical history, as well as the properties of the polypeptide (such as the half-life of the polypeptide, and other considerations well known in the art).

For the purpose of the present disclosure, the appropriate dosage of a therapeutic agent as described herein will depend on the specific agent (or compositions thereof) employed, the formulation and route of administration, the type and severity of the disease, whether the polypeptide is administered for preventive or therapeutic purposes, previous therapy, the subject's clinical history and response to the antagonist, and the discretion of the attending physician. Typically the clinician will administer a polypeptide until a dosage is reached that achieves the desired result.

Administration of one or more polypeptides can be continuous or intermittent, depending, for example, upon the recipient's physiological condition, whether the purpose of the administration is therapeutic or prophylactic, and other factors known to skilled practitioners. The administration of a polypeptide may be essentially continuous over a preselected period of time or may be in a series of spaced dose, e.g., either before, during, or after developing a disease. As used herein, the term "treating" refers to the application or administration of a polypeptide or composition including the polypeptide to a subject in need thereof.

"A subject in need thereof, refers to an individual who has a disease, a symptom of the disease, or a predisposition toward the disease, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve, or affect the disease, the symptom of the disease, or the predisposition toward the disease. In some embodiments, the subject has CDI. In some embodiments, the subject has cancer. In some embodiments, the subject is a mammal. In some embodiments, the subject is a non-human primate. In some embodiments, the subject is human. Alleviating a disease includes delaying the development or progression of the disease, or reducing disease severity. Alleviating the disease does not necessarily require curative results.

As used therein, "delaying" the development of a disease means to defer, hinder, slow, retard, stabilize, and/or postpone progression of the disease. This delay can be of varying lengths of time, depending on the history of the disease and/or individuals being treated. A method that "delays" or alleviates the development of a disease, or delays the onset of the
disease, is a method that reduces probability of developing one or more symptoms of the
disease in a given time frame and/or reduces extent of the symptoms in a given time frame,
when compared to not using the method. Such comparisons are typically based on clinical
studies, using a number of subjects sufficient to give a statistically significant result.

"Development" or "progression" of a disease means initial manifestations and/or
ensuing progression of the disease. Development of the disease can be detectable and assessed
using standard clinical techniques as well known in the art. However, development also refers
to progression that may be undetectable. For purpose of this disclosure, development or
progression refers to the biological course of the symptoms. "Development" includes

occurrence, recurrence, and onset.

As used herein "onset" or "occurrence" of a disease includes initial onset and/or
recurrence. Conventional methods, known to those of ordinary skill in the art of medicine, can
be used to administer the isolated polypeptide or pharmaceutical composition to the subject,
depending upon the type of disease to be treated or the site of the disease. This composition
can also be administered via other conventional routes, e.g., administered orally, parenterally,
by inhalation spray, topically, rectally, nasally, buccally, vaginally or via an implanted
reservoir.

The term "parenteral" as used herein includes subcutaneous, intracutaneous,
intravenous, intramuscular, intraarticular, intraarterial, intrasynovial, intrasternal, intrathecal,
intralesional, and intracranial injection or infusion techniques. In addition, it can be
administered to the subject via injectable depot routes of administration such as using 1-, 3-, or
6-month depot injectable or biodegradable materials and methods.

As used herein, a "subject" refers to a human or animal. Usually the animal is a
vertebrate such as a primate, rodent, domestic animal or game animal. Primates include
chimpanzees, cynomologous monkeys, spider monkeys, and macaques, e.g., Rhesus. Rodents
include mice, rats, woodchucks, ferrets, rabbits and hamsters. Domestic and game animals
include cows, horses, pigs, deer, bison, buffalo, feline species, e.g., domestic cat, canine
species, e.g., dog, fox, wolf, avian species, e.g., chicken, emu, ostrich, and fish, e.g., trout,
catfish and salmon. Patient or subject includes any subset of the foregoing, e.g., all of the
above, but excluding one or more groups or species such as humans, primates or rodents. In
certain embodiments of the aspects described herein, the subject is a mammal, e.g., a primate,

*note: The original text contains a typographical error in "cynomologous", which is clearly intended to be "cynomolgous".

*e.g., a human.
The terms, "patient" and "subject" are used interchangeably herein. A subject can be male or female. A subject can be a fully developed subject (e.g., an adult) or a subject undergoing the developmental process (e.g., a child, infant or fetus). Preferably, the subject is a mammal. The mammal can be a human, non-human primate, mouse, rat, dog, cat, horse, or cow, but are not limited to these examples. Mammals other than humans can be advantageously used as subjects that represent animal models of disorders associated with unwanted neuronal activity. In addition, the methods and compositions described herein can be used to treat domesticated animals and/or pets.

The following examples are intended to be illustrative of certain embodiments and are non-limiting. The entire contents of all of the references (including literature references, issued patents, published patent applications, and co pending patent applications) cited throughout this application are hereby expressly incorporated by reference.

**EXAMPLES**

**Table 1 BoNT Polypeptide Sequences**

<table>
<thead>
<tr>
<th>SEQ ID NO.</th>
<th>Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>WT BoNT/X</td>
<td>MKEIEKFNYPDIDGINVITMRPRHSDKINKEKGPFFGAFQVIKNWIVPER</td>
</tr>
<tr>
<td></td>
<td></td>
<td>YNFTNNTNDLPSEPIMEADAYPNYLNPSEKDFLQGVKLKERIKSKP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EGEKLLELISSLISPLSLSMGATLSDNETITAYQENNNNIVSNLQANLVIYGPG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DIANNATYGLYSTPISENGTGLSEIVSIPEYLFKPDSEGYRNSLVLNIVNFKV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KREFAPDPASTLMHELHYVTHYNLYGISHNRFYNYFD MGMKTEIERTSROQQNSLIFE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ELLETFGGDDKSAISSLIIKJEIATKNNYTLISERLNTVTNEELLYKNNKFVQKPGLGNFKEKTAEFKNLNTIFLVNESNLQROPSILVRKHYLKEREPDIY</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VNILDNYSSTLEFLGNNNSSGNDQGQLLESSYFEKIESNALRAFUKICPRNG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LLYNAIYROSNKLYNNLIDELKKCTSTKNTVSNYPCSLNLGCIEVENKLDFLISN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KSLSNLDINLSEEIKKIPETYFDFKDNLPQDLTSLNYDFTEANSIPSISQNLIER</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NEELYEPIRNSLFTIYVDKLTTFHFEAQNIDESIDSSKRIVELTDVSDEAL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SNPNKVYSPFKNMNTNISIETGTISTYIFQYWLSIRKDFSDETKGIDVIDKSD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SDTLAVIPYIGPLLIGNIDIRHDGFVAIELGITALALLEYVPEFTIPILVGEVIL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGELAREQVEAIVNNDLRDKQWAEVYNITKAQWWGTIHLQINTRLAHT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>YKALSRQANAIKMNMEFQLANYKGNDKAKIKNAISETILLLNSVEQAM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KNTKFKMKSLSLNETKEMIPKQDNKLNFDELITKKTLDKFIKEKEDILGTN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LSSSLRRKVSIRLKNIAFINDIFPSFDDLINQYKNEIYEVNLNGADCEDK</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IKDLSTGTDSDIGIELDADGRENKAIKIKGSEMNIKIIJAMNKYLRFPSATDNF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SISFWIHKPKTPNLLNGIEYTVLENFQVRGWKISIQSDKLIWYLRDHNNSIK</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IVTPDYIAFGNWQIITITNNNKSGIYVYNGSKEEKDDISIWNTEDVPDDPIRI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KNRRNQGDQTLDDQKRIELKNNQENVKLYNNYFNSNYIRDWNQPLQYN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KKYQYQTDQKPGKQLIEYWSFGYDYYVILSDKKTITFPNNIRYGALOYNKS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VLIKNSKQDGLVRKNDQFVLEDGVYMNIGSADVRFNEDTYNGTTYGTGTHDEL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TTDIFIEIQFRKQYKVRNYQOCLKTPYNHKSGLMSTTSKPTFHDYRDWVYSSA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WYFQYNYENLRLNECNWITFYIPKDEGWDED</td>
</tr>
</tbody>
</table>

2          | WT BoNT/X      | MKEIEKFNYPDIDGINVITMRPRHSDKINKEKGPFFGAFQVIKNWIVPER                     |
|            | LC-HN         | YNFTNNTNDLPSEPIMEADAYPNYLNPSEKDFLQGVKLKERIKSKP                          |
KDSLNDINLSEEKIKPETTVFFKDKLPPQDITLSNYDFTEANSIPSISQQNILER
NEELYEPRNLSLEIKITVYDKLTTTFHELAAQNIDESIDSDSKKIRVTLDSDVEAL
SNPNKVYSPFKNMSNTINSIETGIITSTITYFYFQLWRSLIVKDFSDTEGKIDVIDKS
SDTLAVYPGLNLGNIINDRHDGFVGAIELAGITALLEYVEPTFVILGVELIGGELAREQVEAVINNALDKRDQKWAEVYNYTITAQQWGTIHQLINTRAHT
YKALSQANAIKMNMEFQFLANYKGNIDTEDAKAIKNAISSETEILNKSVEQAM
KNTKEFMKLSNLYLTKEMPIKFQVQDNKLNFDELETKTLKDKIKEKEDILGTN
LSSSLRKKVIRLKNIAFDINPFIPEFDDLQNYKNEIEDYVLNGLAGEDGK
IKDLSQTSINDSNGEIDAGRENKAIKIKGSENSTIKIAMNKLYRSATDNF
SIFSWIHKIKPTNLLNGNIEYTLVLVENFQRGWIKISQDSKWLHYRDHNSIKI
VTPDYIAFNGWNLITITNNRSKGSIVYNGSKIEEKDISSSWTEVDDPIRFL
KNNRDQTQATFLDLQPSYRKLDNQNEVVKLYYFNNSYIRDINFGNLPQY
KYYLQTDQKPKGILREYWSFFSYGYVLSDKSTITPFNNIRGALYNQGSK
VLIKNSKDLGVRNKDFEIQLPEIDGYNMGISADRFNEDNTNYGITTGTHTHD
TTDFEIQOKQKYRNYCQLKTPYNIVHKSLMSMTSTKPFTHDYRWDYSSA
WYFYQNYEINLRLKHTKNTWYIPFKDEGWDED

6
WT BoNT/X C467S
MKLEIKNFYNPDIDPNGVNTIMPRPHSRDSKINGKKGPKFKAFAVIKNIWIVPYER
YNFTNNTNLDNIPSEMEIDAINYPNYLNTPSEKDIEFLQVIKVLERIKSKP
EGEDKELIISSSIPPLVSNGALTLDSDNIAYQGENSNNIVSLQLNYLIVYPPG
DIANATYGLYSTPSINGRTLSEVSFSPFYLKPFDSYNGYRSLNVNKFV
KREAPAPSTLMHELHLYLGYSNRNYFYNNDFGTQKIESRQGNSLIFE
ELLTFGGDSKAISSLHIKIKETAKNNYTLISERLTNTVENDKLLYKIN
VQVRLGKPNLTDIAFKEKLNLTIVALNESLIAQRFSILVRKLYKERIPDYI
VNNLDDNSYESTLEGFINSQGSNDQFGQLLSSYFKEKINESALAIKCRPNR
LLYNAIYRNKNLNNLIDKKTSTKNTVSYPCSSLNLGJIEVENKDLFLISN
KDSLNDINLSEEKIKPETTVFFKDKLPPQDITLSNYDFTEANSIPSISQQNILER
NEELYEPRNLSLEIKITVYDKLTTTFHELAAQNIDESIDSDSKKIRVTLDSDVEAL
SNPNKVYSPFKNMSNTINSIETGIITSTITYFYFQLWRSLIVKDFSDTEGKIDVIDKS
SDTLAVYPGLNLGNIINDRHDGFVGAIELAGITALLEYVEPTFVILGVELIGGELAREQVEAVINNALDKRDQKWAEVYNYTITAQQWGTIHQLINTRAHT
YKALSQANAIKMNMEFQFLANYKGNIDTEDAKAIKNAISSETEILNKSVEQAM
KNTKEFMKLSNLYLTKEMPIKFQVQDNKLNFDELETKTLKDKIKEKEDILGTN
LSSSLRKKVIRLKNIAFDINPFIPEFDDLQNYKNEIEDYVLNGLAGEDGK
IKDLSQTSINDSNGEIDAGRENKAIKIKGSENSTIKIAMNKLYRSATDNF
SIFSWIHKIKPTNLLNGNIEYTLVLVENFQRGWIKISQDSKWLHYRDHNSIKI
VTPDYIAFNGWNLITITNNRSKGSIVYNGSKIEEKDISSSWTEVDDPIRFL
KNNRDQTQATFLDLQPSYRKLDNQNEVVKLYYFNNSYIRDINFGNLPQY
KYYLQTDQKPKGILREYWSFFSYGYVLSDKSTITPFNNIRGALYNQGSK
VLIKNSKDLGVRNKDFEIQLPEIDGYNMGISADRFNEDNTNYGITTGTHTHD
TTDFEIQOKQKYRNYCQLKTPYNIVHKSLMSMTSTKPFTHDYRWDYSSA
WYFYQNYEINLRLKHTKNTWYIPFKDEGWDED

7
WT BoNT/X C467A
MKLEIKNFYNPDIDPNGVNTMRPRHSKINGKKGPKFKAFAVIKNIWIVPER
YNFTNNTNLDNIPSEMEIDAINYPNYLNTPSEKDIEFLQVIKVLERIKSKP
EGEDKELIISSSIPPLVSNGALTLDSDNIAYQGENSNNIVSLQLNYLIVYPPG
DIANATYGLYSTPSINGRTLSEVSFSPFYLKPFDSYNGYRSLNVNKFV
KREAPAPSTLMHELHLYLGYSNRNYFYNNDFGTQKIESRQGNSLIFE
ELLTFGGDSKAISSLHIKIKETAKNNYTLISERLTNTVENDKLLYKIN
VQVRLGKPNLTDIAFKEKLNLTIVALNESLIAQRFSILVRKLYKERIPDYI
VNNLDDNSYESTLEGFINSQGSNDQFGQLLSSYFKEKINESALAIKCRPNR
LLYNAIYRNKNLNNLIDKKTSTKNTVSYPCSSLNLGJIEVENKDLFLISN
KDSLNDINLSEEKIKPETTVFFKDKLPPQDITLSNYDFTEANSIPSISQQNILER
NEELYEPRNLSLEIKITVYDKLTTTFHELAAQNIDESIDSDSKKIRVTLDSDVEAL
SNPNKVYSPFKNMSNTINSIETGIITSTITYFYFQLWRSLIVKDFSDTEGKIDVIDKS
SDTLAVYPGLNLGNIINDRHDGFVGAIELAGITALLEYVEPTFVILGVELIGGELAREQVEAVINNALDKRDQKWAEVYNYTITAQQWGTIHQLINTRAHT
YKALSQANAIKMNMEFQFLANYKGNIDTEDAKAIKNAISSETEILNKSVEQAM
KNTKEFMKLSNLYLTKEMPIKFQVQDNKLNFDELETKTLKDKIKEKEDILGTN
LSSSLRKKVIRLKNIAFDINPFIPEFDDLQNYKNEIEDYVLNGLAGEDGK
IKDLSQTSINDSNGEIDAGRENKAIKIKGSENSTIKIAMNKLYRSATDNF
SISFWIKHPKPTNLLNNGIEYTLVENFNQRGWKISIQDSKLIWYLDDHNSIKI
VTPTDYAIGNFNGWNLWLTITTNRNKSGIVSYNGVSKIEEKMISIWNEVDPIFFRFL
KNRRRTDQAFTTLLDQFSYIRKELQNQEVVLKLYNYFNSNYIRDIWGNPLQYN
KKYLYQTQDKPKGRLREYWSFSYGFYDVILSDLSTTTTPFNNIRYGALYNSGK
VLIKNKSDLLRNKVRDFQIDOLEDQYGNMISADRFNEDTNYYGTTYGTTHDL
TTDFEIQQEQEXKYNCQLKTYPNIFHKSGLMSTETSKPTHYDHDWRVDYSSA
WYFQNYENLRKHTKTNYYIFPIKDEGWDDED

8 WT BoNT/X C1240S
MKLEINKFNYNDPIGNVITMRPRPHDSKINKGKGPhFQVKIKNWIIVPER
YNTFTNTNDLNPSEPIMEMADAINYNNLNTPSKEDELFQVGVIKLERISKP
EGEKELLESISPIPLTVSNGALTLDNNETIAYQENNNIVSNQALNYVIGGP
DIANATTYGLYSTPSNGEITLSEVSFSFYKPFDESYGNYSLNIVNKVF
KREAPDPAMTLHELHVTHNLGYGSRNFYYFNDTQGIETSRQNSLIFY
ELLTFQGIDSKAISLIIKQIEAKNNYTTLISERLNTVTENDDLLYYKNIK
VQGLRNKFLIDTDFFKLTLLNLFVESNLQLRFSVLRKHYLERPIDPY
VNILDYNSSTLEGFNNISSQGSNDFOQGQLLESSYFEKIESNALRAFIKCPNRG
LLYNAFVRNSKYNLLNIDLEDKTTSTKNTSYSPCSLLNGCIEVENKDLFLISN
KDSLNDINLSEEKIPETTFFKDLPQDITLNSYDFTEANSISPISQNNLR
NEELEYPIRNLSEIKTIYVDKTLTTFHLEAQNIDESISDSSKIRVETLTDSEAL
SNPNKVSYSPFKMSNTISNITGTSITYFYQWLRIVKDFSDETGKIDVDS
SDLTAVYPIGPLLNLGDIFRDGDFVAIEAGITALLEYVEFTIPILVGEV
GGELEQREVEAIUNNAFLDKRDQKWAENYITKQAQQWGTIIQINTLRALT
YKALSRQANAIAKMNMFQOLNYKIDDDAKAIKNAISETEILNLNSQEVAM
KNTFKMILKNSYTLKEMPQVDNLKNFDLTEKTLDFKEKEGIDLTG
LSSSLRRKVSIIRLNKIAFDINDIPSFEDDLNIQKNEYIDEYVLNLGAEDKG
IKDLSGTTDSIDNGSIEIADGRENKAIKIKSERSTIKAMNYLKRFSATDNF
SISFWIKHPKPTNLLNNGIEYTLVENFNQRGWKISIQDSKLIWYLDDHNSIKI
VTPTDYAIGNFNGWNLWLTITTNRNKSGIVSYNGVSKIEEKMISIWNEVDPIFFRFL
KNRRRTDQAFTTLLDQFSYIRKELQNQEVVLKLYNYFNSNYIRDIWGNPLQYN
KKYLYQTQDKPKGRLREYWSFSYGFYDVILSDLSTTTTPFNNIRYGALYNSGK
VLIKNKSDLLRNKVRDFQIDOLEDQYGNMISADRFNEDTNYYGTTYGTTHDL
TTDFEIQQEQEXKYNCQLKTYPNIFHKSGLMSTETSKPTHYDHDWRVDYSSA
WYFQNYENLRKHTKTNYYIFPIKDEGWDDED
VL1KNSKLDGLVRNKDFIQLEIDGYNMGISADFRENĐTNYIQTGYTTGDHLD TTDFEIQIQREYRNYQNYQLKTPYNIHFHKSGLMSTETSKPTFHDYRDWWYSSA WYFPQNYENLNLRKHTKWYFIPKDEGWDE
NEELYEPIRNSLFEIKTIYVDKLTTFHFLEAQNIDESIDSSKIRVELTDSVDEAL
SNPKVYSPFKNSMNTSINSiETGSITSYIFYQYWLRSIKDFSEDGTKIDVIDKS
SDTLAIVYPIGPLLNIIGNDIRHGDFVAILGILATALLEYYPEFTIPILVGLEI
GELAREQVEAIVNNAKDRLKQKWAEVYNIIAQWGTIHQLINQTRLAHT
YKALSQANAIIKMNMEFQLANYKNIDDKAIKINAISEETLLINNSQVEQAM
KNTKEFKMILSNYLTKEKMPQVDNLKNFDLETKTLDKPKIKEKDELGTN
LSSRLRKVSRILKNNIAFDINIDPSEFDDFLIQ_YKNEI

WT BoNT/X
LC-H5
C461S
MKLEINKFYNDIPIDGINVITMRPRHSDKINKGKGPKFKAQVFKNNIVIFPER
YNTFNNTDLNPISEPMEAIDAINYPNLYNTSEKDEFLQGVKFLKREIKSKP
EGEKLLEISSISLPLVSNGALTLSDNETIAQYQNNINSLQANLYVGPGP
DIANNAYGLYSTPISNGEGLTSVEVSFPLYKDFPDESGINYRSLVNINVFVK
KREFAPDPASTLHELHVTINLGYISNRRNNFYYNFDGTGIETSRQSQNSLIFE
ELLTFGGIDSKAISSLIIKIIETAKNYYTLISERLNVENDLKYIKNKIP
VGQLNFKDRTAEFKKNTLFLVLESN/QRFISLVRKHLYKERPIQDIY
VNLIDNSYSSTLEGPNNSQGNSDFQGQLLESSYEKFIESNALRAFIKCRPRNG
LLYAIYRNSKNYLNINILEDKKSNTKVNSTYPSSSLNCIEVENKDLLFLSLIN
KSDLSNIDNLEEIKPETTVEFSDKLPPQDITLSYNDELETEPISQSOQILNER
NEELYEPIRNSLFEIKTIYVDKLTTFHFLEAQNIDESIDSSKIRVELTDSVDEAL
SNPKVYSPFKNSMNTSINSiETGSITSYIFYQYWLRSIKDFSEDGTKIDVIDKS
SDTLAIVYPIGPLLNIIGNDIRHGDFVAILGILATALLEYYPEFTIPILVGLEI
GELAREQVEAIVNNAKDRLKQKWAEVYNIIAQWGTIHQLINQTRLAHT
YKALSQANAIIKMNMEFQLANYKNIDDKAIKINAISEETLLINNSQVEQAM
KNTKEFKMILSNYLTKEKMPQVDNLKNFDLETKTLDKPKIKEKDELGTN
LSSRLRKVSRILKNNIAFDINIDPSEFDDFLIQ_YKNEI

WT BoNT/X
LC-H5
C467A
MKLEINKFYNDIPIDGINVITMRPRHSDKINKGKGPKFKAQVFKNNIVIFPER
YNTFNNTDLNPISEPMEAIDAINYPNLYNTSEKDEFLQGVKFLKREIKSKP
EGEKLLEISSISLPLVSNGALTLSDNETIAQYQNNINSLQANLYVGPGP
DIANNAYGLYSTPISNGEGLTSVEVSFPLYKDFPDESGINYRSLVNINVFVK
KREFAPDPASTLHELHVTINLGYISNRRNNFYYNFDGTGIETSRQSQNSLIFE
ELLTFGGIDSKAISSLIIKIIETAKNYYTLISERLNVENDLKYIKNKIP
VGQLNFKDRTAEFKKNTLFLVLESN/QRFISLVRKHLYKERPIQDIY
VNLIDNSYSSTLEGPNNSQGNSDFQGQLLESSYEKFIESNALRAFIKCRPRNG
LLYAIYRNSKNYLNINILEDKKSNTKVNSTYPSSSLNCIEVENKDLLFLSLIN
KSDLSNIDNLEEIKPETTVEFSDKLPPQDITLSYNDELETEPISQSOQILNER
NEELYEPIRNSLFEIKTIYVDKLTTFHFLEAQNIDESIDSSKIRVELTDSVDEAL
SNPKVYSPFKNSMNTSINSiETGSITSYIFYQYWLRSIKDFSEDGTKIDVIDKS
SDTLAIVYPIGPLLNIIGNDIRHGDFVAILGILATALLEYYPEFTIPILVGLEI
GELAREQVEAIVNNAKDRLKQKWAEVYNIIAQWGTIHQLINQTRLAHT
YKALSQANAIIKMNMEFQLANYKNIDDKAIKINAISEETLLINNSQVEQAM
KNTKEFKMILSNYLTKEKMPQVDNLKNFDLETKTLDKPKIKEKDELGTN
LSSRLRKVSRILKNNIAFDINIDPSEFDDFLIQ_YKNEI

WT BoNT/X
LC-H5
C467S
MKLEINKFYNDIPIDGINVITMRPRHSDKINKGKGPKFKAQVFKNNIVIFPER
YNTFNNTDLNPISEPMEAIDAINYPNLYNTSEKDEFLQGVKFLKREIKSKP
EGEKLLEISSISLPLVSNGALTLSDNETIAQYQNNINSLQANLYVGPGP
DIANNAYGLYSTPISNGEGLTSVEVSFPLYKDFPDESGINYRSLVNINVFVK
KREFAPDPASTLHELHVTINLGYISNRRNNFYYNFDGTGIETSRQSQNSLIFE
ELLTFGGIDSKAISSLIIKIIETAKNYYTLISERLNVENDLKYIKNKIP
VGQLNFKDRTAEFKKNTLFLVLESN/QRFISLVRKHLYKERPIQDIY
VNLIDNSYSSTLEGPNNSQGNSDFQGQLLESSYEKFIESNALRAFIKCRPRNG
LLYAIYRNSKNYLNINILEDKKSNTKVNSTYPSSSLNCIEVENKDLLFLSLIN
KSDLSNIDNLEEIKPETTVEFSDKLPPQDITLSYNDELETEPISQSOQILNER
NEELYEPIRNSLFEIKTIYVDKLTTFHFLEAQNIDESIDSSKIRVELTDSVDEAL
SNPKVYSPFKNSMNTSINSiETGSITSYIFYQYWLRSIKDFSEDGTKIDVIDKS
SDTLAIVYPIGPLLNIIGNDIRHGDFVAILGILATALLEYYPEFTIPILVGLEI
GELAREQVEAIVNNAKDRLKQKWAEVYNIIAQWGTIHQLINQTRLAHT
YKALSQANAIIKMNMEFQLANYKNIDDKAIKINAISEETLLINNSQVEQAM
KNTKEFKMILSNYLTKEKMPQVDNLKNFDLETKTLDKPKIKEKDELGTN
LSSRLRKVSRILKNNIAFDINIDPSEFDDFLIQ_YKNEI

WT BoNT/X
LC-H5
C467S
MKLEINKFYNDIPIDGINVITMRPRHSDKINKGKGPKFKAQVFKNNIVIFPER
YNTFNNTDLNPISEPMEAIDAINYPNLYNTSEKDEFLQGVKFLKREIKSKP
EGEKLLEISSISLPLVSNGALTLSDNETIAQYQNNINSLQANLYVGPGP
DIANNAYGLYSTPISNGEGLTSVEVSFPLYKDFPDESGINYRSLVNINVFVK
KREFAPDPASTLHELHVTINLGYISNRRNNFYYNFDGTGIETSRQSQNSLIFE
ELLTFGGIDSKAISSLIIKIIETAKNYYTLISERLNVENDLKYIKNKIP
VGQLNFKDRTAEFKKNTLFLVLESN/QRFISLVRKHLYKERPIQDIY
VNLIDNSYSSTLEGPNNSQGNSDFQGQLLESSYEKFIESNALRAFIKCRPRNG
LLYAIYRNSKNYLNINILEDKKSNTKVNSTYPSSSLNCIEVENKDLLFLSLIN
KSDLSNIDNLEEIKPETTVEFSDKLPPQDITLSYNDELETEPISQSOQILNER
NEELYEPIRNSLFEIKTIYVDKLTTFHFLEAQNIDESIDSSKIRVELTDSVDEAL
SNPKVYSPFKNSMNTSINSiETGSITSYIFYQYWLRSIKDFSEDGTKIDVIDKS
SDTLAIVYPIGPLLNIIGNDIRHGDFVAILGILATALLEYYPEFTIPILVGLEI
GELAREQVEAIVNNAKDRLKQKWAEVYNIIAQWGTIHQLINQTRLAHT
YKALSQANAIIKMNMEFQLANYKNIDDKAIKINAISEETLLINNSQVEQAM
KNTKEFKMILSNYLTKEKMPQVDNLKNFDLETKTLDKPKIKEKDELGTN
LSSRLRKVSRILKNNIAFDINIDPSEFDDFLIQ_YKNEI

BoNT/X-LC:
MKLEINKFYNDIPIDGINVITMRPRHSDKINKGKGPKFKAQVFKNNIVIFPER

<table>
<thead>
<tr>
<th>H&lt;sub&gt;N&lt;/sub&gt;-Al-Hc</th>
</tr>
</thead>
<tbody>
<tr>
<td>YNFNTNNTDLNISPSEIPIMEADEAYPNYNLNTPEKDEFQGVIKVELRISKPK</td>
</tr>
<tr>
<td>EGEKELLESSISIPIPLVSNGLTLSNDTIAYQENNNIVSNLQANLVIYVGGP</td>
</tr>
<tr>
<td>DIANATATYGLYSTPISNGEGLTSEVSFPFYKLFDESYGNYRLSNIVNKVF</td>
</tr>
<tr>
<td>KREFAPDPASTLMLVHVTNLHGYSNRNFYYNTDFGLETISQNSLIEF</td>
</tr>
<tr>
<td>ELLTFFGIDDAISSLLHIIKETAKNYNNTTLISERLNTVTVENDLLYKHKIP</td>
</tr>
<tr>
<td>VQGLRGNFKLNTLFVLNESNLQARFISLVVRKHYLKERPIKY</td>
</tr>
<tr>
<td>VNILDNNSYSLFEGNISQGNSDFQGQLLSESYFEKIESALAFIKCRNG</td>
</tr>
<tr>
<td>LLNYAIYRSKNYLNLDDEKKTSKTTSNYPSCSLNGCIEVENKDLFLISN</td>
</tr>
<tr>
<td>KDSDLNDINLEKEIPETTVFSDKLPPQDITLSNYDFTEANSIPSIQNLER</td>
</tr>
<tr>
<td>NEELEYEPIRNLSFEKITYVVDKLTFFLEAQNIDESISKVIRLETSVDEAL</td>
</tr>
<tr>
<td>SNPNSVYSPFKMNSNTINSIEGHTSITYFYFWLRLSIVKDFDSATEGKIDVIDS</td>
</tr>
<tr>
<td>YYTNSVYPIYGPLLNIGNDRHDGVLFAIELGIALLEYVEPTIPVGLGE</td>
</tr>
<tr>
<td>GGE displeQVEIAVNNALDDEKRDQKWAVEYNYITAKQQWWGTTHLQINTLRAHT</td>
</tr>
<tr>
<td>YKALSQRAANAIFMNMQFOLANYGKNDKAKAIKNAISETEILNKSVEQAM</td>
</tr>
<tr>
<td>KNTEKFLMIKSLYNSTLKEMPVQDNKLFNLKDTEKTLDFIKEKEDILGTN</td>
</tr>
<tr>
<td>LSSSLRLRKYSLRKLIAIFDINDIFSFEDDLINQYKNEIIINTSLNRYSENL</td>
</tr>
<tr>
<td>ILDSLYASKINGSKEFYPDFNQQLFLENLSKIEVLNAYINYMENF</td>
</tr>
<tr>
<td>TSWIFRIPYQFNLNLYTNTNEMMNSGWKLSVLNEGYIWTQLITEQIKQ</td>
</tr>
<tr>
<td>RRVFFIKYQSNIADYNRWFIVTITTNNLNSKUNYINGIRDQKPSLNGNIHNA</td>
</tr>
<tr>
<td>SNNMIFLKDCHRTDRYIVKFYNFLDFKELNEEKIDLDYNQNSGILKDFW</td>
</tr>
<tr>
<td>GDYQVDKYPPYMLYLPDNKVDVVNVYIRGMYLKYLPRGSVMVTNNYL</td>
</tr>
<tr>
<td>NSSLRYTGHFKKYASGNKIDVDNVNRDNYIVVINVKKENCYRLATNASQAG</td>
</tr>
<tr>
<td>VEKISLAEIPDVGNSLQGVYMKSNDQGITNNKLMNQDNGNGDFIGF</td>
</tr>
<tr>
<td>HQFPNIALVNSYNIQERSSRTLGCSWEFIPVDGWGGERPL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>23</th>
<th>BoNT/X-LC-H&lt;sub&gt;N&lt;/sub&gt;-B-l-Hc</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKLEINKFNYNPDPIDGINVITMPHRPSDKINGKKGKPKAFAQWIKWIVPER</td>
<td></td>
</tr>
<tr>
<td>YNFNTNNDLNIPEMPEADEAYPNYNLNTPEKDEFQGVIKVELRISKPK</td>
<td></td>
</tr>
<tr>
<td>EGEKELLESSISIPIPLVSNGLTLSNDTIAYQENNNIVSNLQANLVIYVGGP</td>
<td></td>
</tr>
<tr>
<td>DIANATATYGLYSTPISNGEGLTSEVSFPFYKLFDESYGNYRLSNIVNKVF</td>
<td></td>
</tr>
<tr>
<td>KREFAPDPASTLMLVHVTNLHGYSNRNFYYNTDFGLETISQNSLIEF</td>
<td></td>
</tr>
<tr>
<td>ELLTFFGIDDAISSLLHIIKETAKNYNNTTLISERLNTVTVENDLLYKHKIP</td>
<td></td>
</tr>
<tr>
<td>VQGLRGNFKLNTLFVLNESNLQARFISLVVRKHYLKERPIKY</td>
<td></td>
</tr>
<tr>
<td>VNILDNNSYSLFEGNISQGNSDFQGQLLSESYFEKIESALAFIKCRNG</td>
<td></td>
</tr>
<tr>
<td>LLNYAIYRSKNYLNLDDEKKTSKTTSNYPSCSLNGCIEVENKDLFLISN</td>
<td></td>
</tr>
<tr>
<td>KDSDLNDINLEKEIPETTVFSDKLPPQDITLSNYDFTEANSIPSIQNLER</td>
<td></td>
</tr>
<tr>
<td>NEELEYEPIRNLSFEKITYVVDKLTFFLEAQNIDESISKVIRLETSVDEAL</td>
<td></td>
</tr>
<tr>
<td>SNPNSVYSPFKMNSNTINSIEGHTSITYFYFWLRLSIVKDFDSATEGKIDVIDS</td>
<td></td>
</tr>
<tr>
<td>YYTNSVYPIYGPLLNIGNDRHDGVLFAIELGIALLEYVEPTIPVGLGE</td>
<td></td>
</tr>
<tr>
<td>GGE displeQVEIAVNNALDDEKRDQKWAVEYNYITAKQQWWGTTHLQINTLRAHT</td>
<td></td>
</tr>
<tr>
<td>YKALSQRAANAIFMNMQFOLANYGKNDKAKAIKNAISETEILNKSVEQAM</td>
<td></td>
</tr>
<tr>
<td>KNTEKFLMIKSLYNSTLKEMPVQDNKLFNLKDTEKTLDFIKEKEDILGTN</td>
<td></td>
</tr>
<tr>
<td>LSSSLRLRKYSLRKLIAIFDINDIFSFEDDLINQYKNEIIINTSLNRYSENL</td>
<td></td>
</tr>
<tr>
<td>ILDSLYASKINGSKEFYPDFNQQLFLENLSKIEVLNAYINYMENF</td>
<td></td>
</tr>
<tr>
<td>TSWIFRIPYQFNLNLYTNTNEMMNSGWKLSVLNEGYIWTQLITEQIKQ</td>
<td></td>
</tr>
<tr>
<td>RRVFFIKYQSNIADYNRWFIVTITTNNLNSKUNYINGIRDQKPSLNGNIHNA</td>
<td></td>
</tr>
<tr>
<td>SNNMIFLKDCHRTDRYIVKFYNFLDFKELNEEKIDLDYNQNSGILKDFW</td>
<td></td>
</tr>
<tr>
<td>GDYQVDKYPPYMLYLPDNKVDVVNVYIRGMYLKYLPRGSVMVTNNYL</td>
<td></td>
</tr>
<tr>
<td>NSSLRYTGHFKKYASGNKIDVDNVNRDNYIVVINVKKENCYRLATNASQAG</td>
<td></td>
</tr>
<tr>
<td>VEKISLAEIPDVGNSLQGVYMKSNDQGITNNKLMNQDNGNGDFIGF</td>
<td></td>
</tr>
<tr>
<td>HQFPNIALVNSYNIQERSSRTLGCSWEFIPVDGWGGERPL</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>24</th>
<th>BoNT/X-LC-H&lt;sub&gt;N&lt;/sub&gt;-C1-Hc</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKLEINKFNYNPDPIDGINVITMPHRPSDKINGKKGKPKAFAQWIKWIVPER</td>
<td></td>
</tr>
<tr>
<td>YNFNTNNDLNIPEMPEADEAYPNYNLNTPEKDEFQGVIKVELRISKPK</td>
<td></td>
</tr>
<tr>
<td>EGEKELLESSISIPIPLVSNGLTLSNDTIAYQENNNIVSNLQANLVIYVGGP</td>
<td></td>
</tr>
<tr>
<td>DIANATATYGLYSTPISNGEGLTSEVSFPFYKLFDESYGNYRLSNIVNKVF</td>
<td></td>
</tr>
<tr>
<td>KREFAPDPASTLMLVHVTNLHGYSNRNFYYNTDFGLETISQNSLIEF</td>
<td></td>
</tr>
<tr>
<td>ELLTFFGIDDAISSLLHIIKETAKNYNNTTLISERLNTVTVENDLLYKHKIP</td>
<td></td>
</tr>
<tr>
<td>VQGLRGNFKLNTLFVLNESNLQARFISLVVRKHYLKERPIKY</td>
<td></td>
</tr>
<tr>
<td>VNILDNNSYSLFEGNISQGNSDFQGQLLSESYFEKIESALAFIKCRNG</td>
<td></td>
</tr>
<tr>
<td>LLNYAIYRSKNYLNLDDEKKTSKTTSNYPSCSLNGCIEVENKDLFLISN</td>
<td></td>
</tr>
<tr>
<td>KDSDLNDINLEKEIPETTVFSDKLPPQDITLSNYDFTEANSIPSIQNLER</td>
<td></td>
</tr>
</tbody>
</table>
NEELYEPIRNSLFEIKTIYVDKLTTFHFLEAQNIDESIDSSKIRVLTDSVDEAL
SNPNKVYSPFKNMSNTINSIETGTTSTYIFYFQWLRSIVKDFSDETGGKIDVIDKS
SDTLAIVPYIGPLLNGNDIRHDFVGAIELAGITALLEYVPEFTIPLVLGVEI
GGELAREQVEAIVNNALDKRDQKWAEVNYITKAQWWGTHLQINTLAHT
YKALSQRQAIAKMNMEFQLANYKNIDDKAIKNAISETEILNLNVSVEQAM
KNTEKFMIKLSNSOLTKEIMPKVQDLKNFDELEKTTLDFIKEEKIDGLTN
LSSSLRRKVSLRNKNIAFDINDIFSEFDLNNQYKNEIINDSILQLSRQKNT
LVDTSGYNAFEVSEEQDVQLNPFFDHFKLSSGDEGDKVIVTQENIVYNSM
YESIFSIFIRWINKNLSNLPYDITSVKNNSSGWIGIISNFLVFTLQKQEDSE
QSNIFSYDIDSNAPYGYKWFVFTVTNNMIGNMKIYINGKIITDIKTELTG1
NSKFTTFEINKIPDTGLITSDSDNSNMRWIFDYIFAKELDGDKNILEDNLSLQY
TVVVDYWGNDLRNYKEYMYVMDYIDNLNYMYANSRIQIVFNTRRNNDNEF
GYKIIIRKIRGNTNIDTRVGRGIDLYFDMTINTKAYNLFMKNETMYADNHST
EDIYAILGRELQKDINNIFQIPMQMNNTNTYYASIQIKSNFNNGENGICSGITY
RFRLGDWYHRNYLVPVTQKGNYASLESTSTHWGVPVSE

25  BoNT/X-LC-H$_{N}$-Al -C461S

MKLEINKFNYPDIDGINNITMRPPRSHDINKKGGPKFAQVKNIWIVPER
YNFTNNTNDLNIPSEPIMEADAINPNYLNTSEKDEFLQGVKVLERIKSKP
EGEKLEL ISSISSPLVSNGLTLSDNETIAAYQENNINSNLANLQYVIEPP
DIANNATYGLYSTPSNGETLSVECPFYPFLKDFSEGYYRSLVINVKVF
KREFAPPDASTLMHELHVTHNYGISRNFFYNYFNTDGIKSETRSQQNSLFE
ELTTFGGIDSAISSLIIKIKETAKNYYTLISERLNTVTENNDLKYKTNKIP
VQGRLGNFKLDTAEFKKTLNLFLNESLAQRPSFLIRVKHYLKERPIDPIY
VNLDDNSYSTLEGNSQNSDFQQGQLLESLFYKEIESALNRAFKPCRG
LGYNAILYRNSKYNNLNIIDLKTTKTSNYPSLNLGClIEVENKDLFLISN
KDSLNDILSSEEKIPETTVFSDKPQDITLSNYDFTEANSIPISQONLERS
NEELYEPIRNSLFEIKTIYVDKLTTFHFLEAQNIDESIDSSKIRVLTDSVDEAL
SNPNKVYSPFKNMSNTINSIETGTTSTYIFYFQWLRSIVKDFSDETGGKIDVIDKS
SDTLAIVPYIGPLLNGNDIRHDFVGAIELAGITALLEYVPEFTIPLVLGVEI
GGELAREQVEAIVNNALDKRDQKWAEVNYITKAQWWGTHLQINTLAHT
YKALSQRQAIAKMNMEFQLANYKNIDDKAIKNAISETEILNLNVSVEQAM
KNTEKFMIKLSNSOLTKEIMPKVQDLKNFDELEKTTLDFIKEEKIDGLTN
LSSSLRRKVSLRNKNIAFDINDIFSEFDLNNQYKNEIINDSILQLSRQKNT
LVDTSGYNAFEVSEEQDVQLNPFFDHFKLSSGDEGDKVIVTQENIVYNSM
YESIFSIFIRWINKNLSNLPYDITSVKNNSSGWIGIISNFLVFTLQKQEDSE
QSNIFSYDIDSNAPYGYKWFVFTVTNNMIGNMKIYINGKIITDIKTELTG1
NSKFTTFEINKIPDTGLITSDSDNSNMRWIFDYIFAKELDGDKNILEDNLSLQY
TVVVDYWGNDLRNYKEYMYVMDYIDNLNYMYANSRIQIVFNTRRNNDNEF
GYKIIIRKIRGNTNIDTRVGRGIDLYFDMTINTKAYNLFMKNETMYADNHST
EDIYAILGRELQKDINNIFQIPMQMNNTNTYYASIQIKSNFNNGENGICSGITY
RFRLGDWYHRNYLVPVTQKGNYASLESTSTHWGVPVSE

26  BoNT/X-LC-H$_{N}$-B1-Hc-C461S

MKLEINKFNYPDIDGINNITMRPPRSHDINKKGGPKFAQVKNIWIVPER
YNFTNNTNDLNIPSEPIMEADAINPNYLNTSEKDEFLQGVKVLERIKSKP
EGEKLEL ISSISSPLVSNGLTLSDNETIAAYQENNINSNLANLQYVIEPP
DIANNATYGLYSTPSNGETLSVECPFYPFLKDFSEGYYRSLVINVKVF
KREFAPPDASTLMHELHVTHNYGISRNFFYNYFNTDGIKSETRSQQNSLFE
ELTTFGGIDSAISSLIIKIKETAKNYYTLISERLNTVTENNDLKYKTNKIP
VQGRLGNFKLDTAEFKKTLNLFLNESLAQRPSFLIRVKHYLKERPIDPIY
VNLDDNSYSTLEGNSQNSDFQQGQLLESLFYKEIESALNRAFKPCRG
LGYNAILYRNSKYNNLNIIDLKTTKTSNYPSLNLGClIEVENKDLFLISN
KDSLNDILSSEEKIPETTVFSDKPQDITLSNYDFTEANSIPISQONLERS
NEELYEPIRNSLFEIKTIYVDKLTTFHFLEAQNIDESIDSSKIRVLTDSVDEAL
SNPNKVYSPFKNMSNTINSIETGTTSTYIFYFQWLRSIVKDFSDETGGKIDVIDKS
SDTLAIVPYIGPLLNGNDIRHDFVGAIELAGITALLEYVPEFTIPLVLGVEI
GGELAREQVEAIVNNALDKRDQKWAEVNYITKAQWWGTHLQINTLAHT
YKALSQRQAIAKMNMEFQLANYKNIDDKAIKNAISETEILNLNVSVEQAM
KNTEKFMIKLSNSOLTKEIMPKVQDLKNFDELEKTTLDFIKEEKIDGLTN
LSSSLRRKVSLRNKNIAFDINDIFSEFDLNNQYKNEIINDSILQLSRQKNT
LVDTSGYNAFEVSEEQDVQLNPFFDHFKLSSGDEGDKVIVTQENIVYNSM
YESIFSIFIRWINKNLSNLPYDITSVKNNSSGWIGIISNFLVFTLQKQEDSE
QSNIFSYDIDSNAPYGYKWFVFTVTNNMIGNMKIYINGKIITDIKTELTG1
NSKFTTFEINKIPDTGLITSDSDNSNMRWIFDYIFAKELDGDKNILEDNLSLQY
TVVVDYWGNDLRNYKEYMYVMDYIDNLNYMYANSRIQIVFNTRRNNDNEF
GYKIIIRKIRGNTNIDTRVGRGIDLYFDMTINTKAYNLFMKNETMYADNHST
EDIYAILGRELQKDINNIFQIPMQMNNTNTYYASIQIKSNFNNGENGICSGITY
RFRLGDWYHRNYLVPVTQKGNYASLESTSTHWGVPVSE
SSLYRGTKFIICKYASGNKDINVRNDRVYINVVVKNEKYRLATNASQGV
EKLISALEIPDVGNLQGQQVVMKSDKNDQGITNKCKMNILQDNNGNDIGFIHF
QFNNIAKLVASNYWRQIERSSRTLCSCWEFIPVDDGWGERPL

29

BoNT/X-LC-
H1-B1-Hc
C467S

MKLEINKFNYNDPDDGINVITMRRPPRHSDDKIKGKPFFAFQVIKINWIVER
YNFTNNTNLDNLIEPEMADAIYPNLNLPSEKDFLGQVICKELIKSKP
EGEKKLEISSSSIPLPLVSNGALTLSDEIAYQENNNVSNLQANLVIYGPGP
DIANNATGYLSTPSEGTELSEVSFSFYLKFDESGNYLRSLNNVIFK
KREFAAPDPASTLMEHLHLVHTHNLYGISNRRNFYNNFDTGKIJERSRQNSLIFEE
LLTFGGIDSKAISLLIKKIEATKNNYTTLISERLNTTVENDDLLKYIKNIPP
VQRGLNFKLDTAEFEKKNLNTILFVLNESLQAFSIILVRKHYLKERPIDFIV
YLILDNSYSLEGNNISQGNSDFQGLLESSYFKEIEKNALAFIKCPNRLG
LYNAIYRNKSYLNNILEDKKTSTKTVSYPCSLLINGSJEVENKDLFJNSK
DLSLNdNLSEEKIPPEPTTVFKDLPPQDITLSYNTDEAANAIPSISQNIKL
EELGEPINLSLEIYTIYDKLIYTFHLEAQIUNIDESDSKIRVLTSDSEALS
PNPKVYSPFKMSNTSINTSITGETSTIFYQYWRLSIVKDFSDTEGTIDKVIDKSS
DTLAIYVPYGIPLLNIGDRHGDNFGVIEGALTALLEYVEPFITVLEIYG
GELAREQVEAIVNNALDKRDQKWAQYVFESLKAQWDHGTLQIVTRLALHTY
KALSQQAIAKGMNMEQFQANYGNDDDIAKNAISEETILNNKSEVAMK
NTEKFMIKLNSYLTKEMKIPVQDNLKNFDELTKKTIKLFKIEKEADMILTGNL
SLRRKVSRLKNNAFDINPSEFDDDLNQYKNEILNNILNNLYRDNNLID
LSGYGAKVEYDGVLKDNQKFGLTSASSVIKTQVNNQINIFNSFVLDFSVS
FWIRIPKYKNGDIQYINHEYTIINCMEKNSWSKIRGNIWTLDINGTK
SVFVEFAYENINREVWNTVLNNLNNITYINGKLESNEDKIDEVRIANG
EIIKFLGDDIDRTQFIWKMKYFISFINTLSQNEEYRKYQSEYLKDFWGNPL
MKNKEYYMFANGNKSYLKLKDSPEVLTSEKRYNKSYNdNYRIILYG
FLKFIIRKNSNSQINDIVKEDY1YDLFNLQERWRYTYKFKKEEKLFLA
PISDSDFYNTIQHKEYDEQTPYSCQALLFKEEDESTEDEGLIGHIYFESGIVF
EYKDYFCISKKWLYKEVKKRPYNLKLGNCWQPFPIDGEWTE

30

BoNT/X-LC-
H1-Cl-Hc
C467S

MKLEINKFNYNDPDDGINVITMRPPRHSDDKIKGKPFFAFQVIKINWIVER
YNFTNNTNLDNLIEPEMADAIYPNLNLPSEKDFLGQVICKELIKSKP
EGEKKLEISSSSIPLPLVSNGALTLSDEIAYQENNNVSNLQANLVIYGPGP
DIANNATGYLSTPSEGTELSEVSFSFYLKFDESGNYLRSLNNVIFK
KREFAAPDPASTLMEHLHLVHTHNLYGISNRRNFYNNFDTGKIJERSRQNSLIFEE
LLTFGGIDSKAISLLIKKIEATKNNYTTLISERLNTTVENDDLLKYIKNIPP
VQRGLNFKLDTAEFEKKNLNTILFVLNESLQAFSIILVRKHYLKERPIDFIV
YLILDNSYSLEGNNISQGNSDFQGLLESSYFKEIEKNALAFIKCPNRLG
LYNAIYRNKSYLNNILEDKKTSTKTVSYPCSLLINGSJEVENKDLFJNSK
DLSLNdNLSEEKIPPEPTTVFKDLPPQDITLSYNTDEAANAIPSISQNIKL
EELGEPINLSLEIYTIYDKLIYTFHLEAQIUNIDESDSKIRVLTSDSEALS
PNPKVYSPFKMSNTSINTSITGETSTIFYQYWRLSIVKDFSDTEGTIDKVIDKSS
DTLAIYVPYGIPLLNIGDRHGDNFGVIEGALTALLEYVEPFITVLEIYG
GELAREQVEAIVNNALDKRDQKWAQYVFESLKAQWDHGTLQIVTRLALHTY
KALSQQAIAKGMNMEQFQANYGNDDDIAKNAISEETILNNKSEVAMK
NTEKFMIKLNSYLTKEMKIPVQDNLKNFDELTKKTIKLFKIEKEADMILTGNL
SLRRKVSRLKNNAFDINPSEFDDDLNQYKNEILNNILNNLYRDNNLID
LSGYGAKVEYDGVLKDNQKFGLTSASSVIKTQVNNQINIFNSFVLDFSVS
FWIRIPKYKNGDIQYINHEYTIINCMEKNSWSKIRGNIWTLDINGTK
SVFVEFAYENINREVWNTVLNNLNNITYINGKLESNEDKIDEVRIANG
EIIKFLGDDIDRTQFIWKMKYFISFINTLSQNEEYRKYQSEYLKDFWGNPL
MKNKEYYMFANGNKSYLKLKDSPEVLTSEKRYNKSYNdNYRIILYG
FLKFIIRKNSNSQINDIVKEDY1YDLFNLQERWRYTYKFKKEEKLFLA
PISDSDFYNTIQHKEYDEQTPYSCQALLFKEEDESTEDEGLIGHIYFESGIVF
EYKDYFCISKKWLYKEVKKRPYNLKLGNCWQPFPIDGEWTE
31 BoNT/X
R360A/Y363 F
MKLEINKFNYPDIDGINVTMRPPRRSDKINGKGPGPKAFQVIKNIVIPER
YNTFTNNTDLNIPSEPMIDAEAIYPNYLNTPSKDEFLQGVIKLIERKSKP
EGEKKLELISSIPPLPSVGLTLSNETIAQENNNLVSNQLVNYIVGVG
DIANNATYGLYSTPSNQEGTSLVSESFPSYLKPDFESYGNRSVLYNVNFKV
KREFAPDPASTMLVHVTNYLGISNRFNYYNFDTGTKIERTSQNSLIFEE
LLTTFGGIDSKAIISSLUIKJEKTNNTILSSLTVENDLKLKRYNKIPV
QGRGLNFKLDTAEFFKLNTLVLVSNLQAFILYEVNLGLAEQGKIKL
DLSGTTSDINSGDIELADGRENKAIKICKANSTIKAKNYLFRSALTDSFISI
FWIKHKPNTNLNNGIETYLVNENFQRGWSIKQDSKLYLWYRLDHNNSIKIIT
PDYIAFNGWNLITITTNNRSKGISYYVNGSKEEKDISSWNTVEDFIIIFRLK
NRDSTDQATFLDLQFSYREKLQNEVNEVKLNYFNSNYRHDWNLQYNKK
YLUYTQDTPKGLIREYWSSFETYVLSDSKTIFPNRNYGALYNGKVLI
KNSKXGLPVKDFQLEIDGGYNMIGADRFEDNTNYIGTGYTHDLTTD
FEIIQFRQKEYKRNQKTYPIFHKSGMLSTETSKFTHYDVRWYSSAWYF
QNYENLNLKRTKWNYPFIPDEGGWDED

32 BoNT/X
H227Y
MKLEINKFNYPDIDGINVTMRPPRRSDKINGKGPGPKAFQVIKNIVIPER
YNTFTNNTDLNIPSEPMIDAEAIYPNYLNTPSKDEFLQGVIKLIERKSKP
EGEKKLELISSIPPLPSVGLTLSNETIAQENNNLVSNQLVNYIVGVG
DIANNATYGLYSTPSNQEGTSLVSESFPSYLKPDFESYGNRSVLYNVNFKV
KREFAPDPASTMLVHVTNYLGISNRFNYYNFDTGTKIERTSQNSLIFEE
LLTTFGGIDSKAIISSLUIKJEKTNNTILSSLTVENDLKLKRYNKIPV
QGRGLNFKLDTAEFFKLNTLVLVSNLQAFILYEVNLGLAEQGKIKL
DLSGTTSDINSGDIELADGRENKAIKICKANSTIKAKNYLFRSALTDSFISI
FWIKHKPNTNLNNGIETYLVNENFQRGWSIKQDSKLYLWYRLDHNNSIKIIT
PDYIAFNGWNLITITTNNRSKGISYYVNGSKEEKDISSWNTVEDFIIIFRLK
NRDSTDQATFLDLQFSYREKLQNEVNEVKLNYFNSNYRHDWNLQYNKK
YLUYTQDTPKGLIREYWSSFETYVLSDSKTIFPNRNYGALYNGKVLI
KNSKXGLPVKDFQLEIDGGYNMIGADRFEDNTNYIGTGYTHDLTTD
FEIIQFRQKEYKRNQKTYPIFHKSGMLSTETSKFTHYDVRWYSSAWYF
QNYENLNLKRTKWNYPFIPDEGGWDED

33 BoNT/X
E228Q
MKLEINKFNYPDIDGINVTMRPPRRSDKINGKGPGPKAFQVIKNIVIPER
YNTFTNNTDLNIPSEPMIDAEAIYPNYLNTPSKDEFLQGVIKLIERKSKP
EGEKKLELISSIPPLPSVGLTLSNETIAQENNNLVSNQLVNYIVGVG
DIANNATYGLYSTPSNQEGTSLVSESFPSYLKPDFESYGNRSVLYNVNFKV
KREFAPDPASTLMHQLVHVTHNLYGISNRNFYYNFDTKIETSRQNSLIFE
ELLTFGCGIDSKAISLLIKKIIETAKNNYTTILSIRELNTVTVERNDLLKYIKNKIPV
QGLRLGNFKLDTELAEFFKLNTLFLVNESNLAQRFISILYKHYLKERPDPYYPV
NILDDNSYSTLEGFSNSGSDFQGLLLESSYFEKIESNALRAFIKCPNRGL
LYNAIYRSNKNYLNNLIDDLKTSKTNVSYPCLSNLGCEVENKDLFSLNK
DSDLINLSEEKIPETTVFFKDLPQDITLSNYDFTEANSISPSIQNILEN
EELYEIPINSFLFEIKITVYDKLTTTFHELAEJQIDESISDSKIRVELTDSIDEALS
NPNKYSVPFKMNSINTSIEGTSTIFYQWLRSIVKDFSDEGTKGIDVKS
DTLAIVPYGYPLNIGNDRHRDFVGAIELAGITALLEYVEFTPIVGLVEIGV
GELAREQVEAIIVNNALDKRDQKWAVYNITAKQWGYTHTLQINTLRAHTY
KALSRSQAIAKMNMEFQALANYKQNIDDKAKIKNAISETEILLNSVEQAMK
NTEKFMIKLSNYLTMEKIMPVQDNLKNFDLETKTDLDFKEKEILDLGNL
SLLRRKYSIRLKNIAFINDIPSEFDLINVKYNEIEDEYVNLGAEAGDKK
DLSGTTSIDINSGDIAEDRGAKIAKIKGSENSTIKIAMNKLRFSAIDNFSIS
FWIKHPKFNTKNMNFISSIPPQDLVSNATLSDNIELAYQENNNIVSNLQLVLYVGPG
DIANNATYGLYSTPSIEGTGEVLSFSFPPYKFDESYGNYRSLVNVNKVFV
KREFAPDAMHNHELVYHTHLYSINFRNNYNFNDTGIETSRQONSLIFEE
LTTFGGIDSKAISLLIKKIIETAKNNYTTILSIRELNTVTVERNDLLKYIKNKIPV
QGLRLGNFKLDTELAEFFKLNTLFLVNESNLAQRFISILYKHYLKERPDPYYPV
NILDDNSYSTLEGFSNSGSDFQGLLLESSYFEKIESNALRAFIKCPNRGL
LYNAIYRSNKNYLNNLIDDLKTSKTNVSYPCLSNLGCEVENKDLFSLNK
DSDLINLSEEKIPETTVFFKDLPQDITLSNYDFTEANSISPSIQNILEN
EELYEIPINSFLFEIKITVYDKLTTTFHELAEJQIDESISDSKIRVELTDSIDEALS
NPNKYSVPFKMNSINTSIEGTSTIFYQWLRSIVKDFSDEGTKGIDVKS
DTLAIVPYGYPLNIGNDRHRDFVGAIELAGITALLEYVEFTPIVGLVEIGV
GELAREQVEAIIVNNALDKRDQKWAVYNITAKQWGYTHTLQINTLRAHTY
KALSRSQAIAKMNMEFQALANYKQNIDDKAKIKNAISETEILLNSVEQAMK
NTEKFMIKLSNYLTMEKIMPVQDNLKNFDLETKTDLDFKEKEILDLGNL
SLLRRKYSIRLKNIAFINDIPSEFDLINVKYNEIEDEYVNLGAEAGDKK
DLSGTTSIDINSGDIAEDRGAKIAKIKGSENSTIKIAMNKLRFSAIDNFSIS
FWIKHPKFNTKNMNFISSIPPQDLVSNATLSDNIELAYQENNNIVSNLQLVLYVGPG
DIANNATYGLYSTPSIEGTGEVLSFSFPPYKFDESYGNYRSLVNVNKVFV
KREFAPDAMHNHELVYHTHLYSINFRNNYNFNDTGIETSRQONSLIFEE
LTTFGGIDSKAISLLIKKIIETAKNNYTTILSIRELNTVTVERNDLLKYIKNKIPV
QGLRLGNFKLDTELAEFFKLNTLFLVNESNLAQRFISILYKHYLKERPDPYYPV
NILDDNSYSTLEGFSNSGSDFQGLLLESSYFEKIESNALRAFIKCPNRGL
LYNAIYRSNKNYLNNLIDDLKTSKTNVSYPCLSNLGCEVENKDLFSLNK

34 BoNT/X
H231Y

MKLEINKFNYYNDPIDIGVINVTMRRPPHSDKINKKGKPGPFKAFQVIKNIWIVPER
YNFTNNTDNLPISPEMADIAINPNYNLTPTSEKDOLQVYKVLERIKSKP
EGEKELILSEISIPPLVSNALTSDNIELAYQENNNIVSNLQLVLYVGPG
DIANNATYGLYSTPSIEGTGEVLSFSFPPYKFDESYGNYRSLVNVNKVFV
KREFAPDAMHNHELVYHTHLYSINFRNNYNFNDTGIETSRQONSLIFEE
LTTFGGIDSKAISLLIKKIIETAKNNYTTILSIRELNTVTVERNDLLKYIKNKIPV
QGLRLGNFKLDTELAEFFKLNTLFLVNESNLAQRFISILYKHYLKERPDPYYPV
NILDDNSYSTLEGFSNSGSDFQGLLLESSYFEKIESNALRAFIKCPNRGL
LYNAIYRSNKNYLNNLIDDLKTSKTNVSYPCLSNLGCEVENKDLFSLNK
DSDLINLSEEKIPETTVFFKDLPQDITLSNYDFTEANSISPSIQNILEN
EELYEIPINSFLFEIKITVYDKLTTTFHELAEJQIDESISDSKIRVELTDSIDEALS
NPNKYSVPFKMNSINTSIEGTSTIFYQWLRSIVKDFSDEGTKGIDVKS
DTLAIVPYGYPLNIGNDRHRDFVGAIELAGITALLEYVEFTPIVGLVEIGV
GELAREQVEAIIVNNALDKRDQKWAVYNITAKQWGYTHTLQINTLRAHTY
KALSRSQAIAKMNMEFQALANYKQNIDDKAKIKNAISETEILLNSVEQAMK
NTEKFMIKLSNYLTMEKIMPVQDNLKNFDLETKTDLDFKEKEILDLGNL
SLLRRKYSIRLKNIAFINDIPSEFDLINVKYNEIEDEYVNLGAEAGDKK
DLSGTTSIDINSGDIAEDRGAKIAKIKGSENSTIKIAMNKLRFSAIDNFSIS
FWIKHPKFNTKNMNFISSIPPQDLVSNATLSDNIELAYQENNNIVSNLQLVLYVGPG
DIANNATYGLYSTPSIEGTGEVLSFSFPPYKFDESYGNYRSLVNVNKVFV
KREFAPDAMHNHELVYHTHLYSINFRNNYNFNDTGIETSRQONSLIFEE
LTTFGGIDSKAISLLIKKIIETAKNNYTTILSIRELNTVTVERNDLLKYIKNKIPV
QGLRLGNFKLDTELAEFFKLNTLFLVNESNLAQRFISILYKHYLKERPDPYYPV
NILDDNSYSTLEGFSNSGSDFQGLLLESSYFEKIESNALRAFIKCPNRGL
LYNAIYRSNKNYLNNLIDDLKTSKTNVSYPCLSNLGCEVENKDLFSLNK

35 BoNT/X-LC-
H N
R360A/Y363 F

MKLEINKFNYYNDPIDIGVINVTMRRPPHSDKINKKGKPGPFKAFQVIKNIWIVPER
YNFTNNTDNLPISPEMADIAINPNYNLTPTSEKDOLQVYKVLERIKSKP
EGEKELILSEISIPPLVSNALTSDNIELAYQENNNIVSNLQLVLYVGPG
DIANNATYGLYSTPSIEGTGEVLSFSFPPYKFDESYGNYRSLVNVNKVFV
KREFAPDAMHNHELVYHTHLYSINFRNNYNFNDTGIETSRQONSLIFEE
LTTFGGIDSKAISLLIKKIIETAKNNYTTILSIRELNTVTVERNDLLKYIKNKIPV
QGLRLGNFKLDTELAEFFKLNTLFLVNESNLAQRFISILYKHYLKERPDPYYPV
NILDDNSYSTLEGFSNSGSDFQGLLLESSYFEKIESNALRAFIKCPNRGL
LYNAIYRSNKNYLNNLIDDLKTSKTNVSYPCLSNLGCEVENKDLFSLNK
GELAREQVEAIVNNALDKRDQKWAEVYNITKAQWWTIHLQINTRALHTYKALSQRANAIKMNFMEQFLANYNGIDDDAKIKNAISETILNNKSVEQAMKNTKIFMIKLSNYTLMKEMPVQDNLKNFDFLTEKTDLKFIKEKEDILGTNLSSLRRKSVILRNKINNIDINDIPSEFDDLINQYKNEI

39 VAMP1  MSAPQAQPPEAGTETAPPGGPPPNNMTSNRRLQQTQAAEVEVVDFDIRVNDKVLERDQKLSLELDRADALQAGASQFESSAAKLRKRKYWWKNCKMMIMMGLAICAIIVVVIVIIFT

40 VAMP2  MSATAATAPPAAPGEGGPPAPPN LTSNRRLLQQTQAQDEVVDMRNVNYDKVLERDQKLSLELDRADALQAGASQFETAAKLRKRKYWWKNKMMILLGVCIAIIIIVIVYFFS

41 VAMP3  MSTGVGVSAAATGNSNRLQQTQVQDEVVDIMMVWDKVLERDQKSLSELDDRADALQAGASQFETAAKLRKRKYWWKNCKMWAIGISVLVIIVIIIIVVCS

42 VAMP4  MPPFKKRHLNDDVTGVSKERRNLLEDDDEEEDDDLRGPSGFPRGPRNDKIKHVQVQDEVDEVMQENITKVIERGERLEDELQKSESLSDNDANAFSNRSKQLRRQMQWWRGCCKIKIAMALAAIILLVILLVIMKRYT

43 VAMP5  MAGIELERQQANVEITEIRMNRFNFGVLRGKLAELQQRSDQDLDMSSTFNTKTTONLQAJKCCWENYRICVGLVVGILLIVLVVVFQPQSSDSAPRTQDAGIAASPGPN

44 Yki6  MKLYLSVLKYGKEAKVVLKAAAYDVSSSFQFQSVQEFMFTSQUALVERS KGTASVKEQDYLHVYRNDLAVSAGVLVIDAYNEYPSKVAFTLKLVELDFSQKVDRIDPVPVSAPTIIAEDLGSLRQYSNPREADPMPWQAEL DLKTHIILHNMESLLEGERKLNDDLVSKEVLGTQSAFYKTARKQNSCCAI

45 BoNT/X-LC-HN-LPETGG  MKLEINKFNYNDPIDGGINVTMRPPRHSDDKINKGKGPFKAFQVQVKNINYPER YNFNTDRNLDNIPSEPMIQUEADIAYNLNTLPSEKDELQFGVYKVLERIKSKPE GEGKLLDELISIPPLVSNGALTLDSENIAQVENNNLSNQLANLYVYGPGPNI DDIATAGLYSTPSINGETGHELSEVSFSFYPFLKPFESYNGYRSLNYVKNFV KREAFPAPDSTMLHVELVHTNLYGINSNRFYNFECDTGIERTSQNSLFEEL LITFEGIDPASILKKIKIITEKANNYTTISSLRENLTVENDYKLKYKINPKV QGRLNGKFDLTAEDKLTLNLVLESNAQLQFSILVRYHLKYLRPDIQYYNV ILDDNSYSTLEGSQGSQDFFQQQOLESYEFIESKIANLAFIKCPRNGL LNYAICYRNSKNYLNINLEDJKTVSTKNTVSYYLPSCSLNGCIEVENKDLFLISNS DLSNDINLSEEKIPETTVFFSDKLPLPQDTILSNYDFTEANSIPISQQNILERN ELEYEPNMLSEIXTIVVDKLTTHFLEAQNIDESDSKIRVLFTSDSEALDS NPKNYVSPFKMNSTINTETGTSYTFQYFWLRSLVSFDSDETGKIDVIDKS DTLAIAPYIGPLNIGNIDHRDFVGAEJAILITALLEVYEPIFTLPILVLEIGGVE GELAREQVEAIVNNALDKRDQKWAEVYNITKAQWWTIHLQINTRALHTYKALSQRANAIKMNFMEQFLANYNGIDDDAKIKNAISETILNNKSVEQAMKNTKIFMIKLSNYTLMKEMPVQDNLKNFDFLTEKTDLKFIKEKEDILGTNLSSLRRKSVILRNKINNIDINDIPSEFDDLINQYKNEI

46 G-BoNT/X-Hc  GEDYVLEVNLGAEDGKDKLSGTTSINDISDIELADGRENKAIKIKGSENSTIK IAMNKYLRDFATDNSFISFWIKHPKTNLNNNGIEYTLVENENQRGWQSIKO SKLFYWRMLHNSLKIVTEDPYAFINGWNLITITNNSRGISYVYGSKIEKDI SSSIWEETDVPPIIFRKLKNRDTQAFKLDDIFSIRYRKLONENVYKLKNYYFN SNYIRDGNPLQYNNKQYLTQDPPKGLREYWSFFGYYDVLISDKTIFFFNPIR YNGALYSKVLKNSKSKDLVNLQFIQLEDQGINNIGSADFRDNTNYGTFFGTTHDLTTDFEIIQERQKEYRMTCQKLTPYNHFKGLSTETSPTPTHYRDITYWSSAWYFYQNYENLNRLKHTKTNWFYPKDEGWD

47 BoNT/Al-Hc  IINTSILNLRYEHSNHLDSLSKLYASIGNIKGSKVFDHDPDKNIQOFNLNESSKIEVL KNAIYVNSYMFNSTSNFWSIFRYPKFYFNSHLNEYTNINCMENNSGWKLYNG EIIWTLQDQETIHYQYRFSQHMINISDYRNFVTITNNLRNSKNPYYINGRL DQPKISPNIHMNDSSLKFMGLDGRCDTHYIWIYFNLFKELNEKEIDA DQNNSNGSILDFGDQLYQYDKPPYMLNILDPKYYVDNVNIGIRGYMLPK PGRSVMNITNNLNSLYRTGIFKIKYASGKNKDINVRNDRYVNVYNNVVKNKE YRALTANASHGQVEKILSALEIPDVGNSLSQVSVYMKSDKNGTITNCMKLQ DNNNGNDIGFIFGDFQFNNIAKLVASNWYNYQIERSSRTLCGERWEFIPVDDWG
**A Novel Botulinum Neurotoxin and Its Derivatives**

Botulinum neurotoxins (BoNTs) are among the most dangerous potential bioterrorism agents and are also used clinically to treat a growing list of medical conditions. There are seven serotypes of BoNTs (BoNT/A-G) known to date, and no new types have been recognized for the past 45 years. Ggenomic database searching of *Clostridium botulinum* strains revealed a novel BoNT type, named BoNT/X. This toxin showed the lowest sequence identity with other BoNTs and it is not recognized by antisera raised against known BoNT types. It cleaves vesicle associated membrane protein (VAMP) in neurons, which is also the target of BoNT/B/D/F/G, but BoNT/X cleaves at a site (between Arg66-Ala67 on VAMP2) unique to this toxin. To validate the activity of BoNT/X, a limited amount of full-length BoNT/X were assembled by covalently linking two non-toxic fragments of BoNT/X using a transpeptidase (sortase). Assembled BoNT/X entered cultured neurons and cleaved VAMP2, and caused flaccid
paralysis in mice measured by Digit Abduction Score assay. Together, these data established BoNT/X as a novel BoNT type. Its discovery poses an urgent challenge for developing effective countermeasures and also presents a novel tool for potential therapeutic applications.

Searching genomic databases revealed a novel BoNT gene

In an attempt to survey the evolutionary landscape of BoNTs, iterative Hidden Markov model searches of the PubMed sequence database were performed, utilizing sequences of the seven BoNTs as probes. The search successfully identified major BoNT serotypes, subtypes, and mosaic toxins, as well as related tetanus neurotoxin (TeNT) (FIG. 5). Unexpectedly, it also revealed a novel BoNT gene (GenBank No. BAQ 12790.1), from the recently reported genomic sequence of Clostridium botulinum strain 111. This toxin gene is herein designated as BoNT/X.

Phylogenetic analysis revealed that BoNT/X is clear distinct from all other BoNTs and TeNT (FIG. 1A). It has the least protein sequence identity (<31%) from any other BoNTs among pair-wise comparisons within BoNT/TeNT family (FIG. 1A). For instance, BoNT/A and BoNT/B share 39% sequence identity, and BoNT/B and BoNT/G have 58% sequence identity. Furthermore, a sliding sequence comparison window demonstrated that the low similarity is evenly distributed along BoNT/X sequence as compared to the other seven BoNTs and TeNT (FIG. IB), indicating that it is not a mosaic toxin.

Despite the low sequence identity, the overall domain arrangement and a few key features of BoNTs appear to be conserved in BoNT/X (FIG. IB), including: (1) a conserved zinc-dependent protease motif HExxH (residues 227-231, HELVH (SEQ ID NO: 92)) is located in the putative LC; (2) there are two conserved cysteines located at the border between the putative LC and HC, which may form the essential inter-chain disulfide bond; (3) a conserved receptor binding motif SxWY exists in the putative He (residues 1274-1277, SAWY (SEQ ID NO: 93)), which recognizes lipid co-receptor gangliosides 4\textsuperscript{344}.

As expected, BoNT/X gene is preceded with a putative NTNHA gene (FIG. 1C). They are located in an OrfX gene cluster. However, the OrfX gene cluster of BoNT/X has two unique features compared to the other two known OrfX clusters (FIG. 1C): (1) there is an additional OrfX2 protein (designated as OrfX2b) located next to the BoNT/X gene, which has not been reported for any other OrfX clusters; (2) the reading frame of OrfX genes has the same direction with the BoNT/X gene, while they are usually opposite to the direction of BoNT gene in other OrfX clusters (FIG. 1C). Together, these features suggest that BoNT/X
may constitute a unique evolutionary branch of the BoNT family.

**The LC of BoNT/X cleaves VAMP2 at a novel site**

Whether BoNT/X is a functional toxin was next examined. First, the LC of BoNT/X (X-LC) was investigated. The border of the LC (residues 1-439) was determined by sequence alignment with other BoNTs. The cDNA encoding the LC was synthesized and the LC was produced as a His6-tagged recombinant protein in *E. coli*. X-LC was incubated with rat brain detergent extracts (BDE) and immunoblot analysis was used to examine whether the three dominant SNARE proteins in the brain, SNAP-25, VAMP2, and syntaxin 1, were cleaved. LCs of BoNT/A (A-LC) and BoNT/B (B-LC) were assayed in parallel as controls. Cleavage of SNAP-25 by BoNT/A generates a smaller fragment that can still be recognized on immunoblot, while cleavage of VAMP2 by BoNT/B abolishes the immunoblot signal of VAMP2 (FIG. 2A). Synaptophysin (Syp), a synaptic vesicle protein, was also detected as an internal loading control. Incubation of X-LC with rat brain DTE did not affect syntaxin 1 or SNAP-25, but abolished VAMP2 signals (FIG. 2A). LCs of BoNTs are zinc-dependent proteases. As expected, EDTA prevented cleavage of SNARE proteins by X-, A-, and B-LCs (FIG. 2A). To further confirm that X-LC cleaves VAMP2, the cytosolic domain of VAMP2 (residues 1-96) as a His6-tagged protein was purified. Incubation of VAMP2 (1-96) with X-LC converted the VAMP2 band into two lower molecular weight bands on SDS-PAGE gel (FIG. 2B), confirming that X-LC cleaves VAMP2.

To identify the cleavage site on VAMP2, the VAMP2 (1-96) protein was analyzed with or without pre-incubation with X-LC, by liquid chromatography-tandem mass spectrometry (LC-MS/MS, FIGs. 2C-2E, see below for detail). A single dominant peptide peak appeared after incubation with X-LC (FIGs. 2C, 2E, and 6). Its molecular weight was determined to be 308.17, which fits only the peptide sequence of residues A67-L96 of VAMP2 (FIGs. 2C, 2E). Consistently, the other fragment from the beginning of the His6-tag to the residue R66 of VAMP2 was also detected (FIG. 2D). To further confirm this result, the assay was repeated with a different VAMP2 fragment: GST-tagged recombinant VAMP2 (33-86) (FIG. 7). Incubation with X-LC generated a single dominant peak, with a molecular weight of 2063.1, which fits only the peptide sequence of residues A67-R86 of VAMP2 (FIGs. 7D-7E). As expected, the other fragment from the beginning of the GST tag to the residue R66 of VAMP2 was also detected (FIG. 7F). Together, these results demonstrated that X-LC has a single cleavage site on VAMP2 between R66 and A67.
R66-A67 is a novel cleavage site distinct from established target sites for all other BoNTs (FIG. 2F). It is also the only BoNT cleavage site located within a region previously known as SNARE motif (FIG. 2F, shaded regions) 45. VAMP family proteins include VAMP1, 2, 3, 4, 5, 7, 8, as well as related Sec22b and Ykt6. R66-A67 is conserved in VAMP1 and VAMP3, which are highly homologous to VAMP2, but not in other VAMP homologs such as VAMP7 and VAMP8. To validate the specificity of X-LC, HA-tagged full-length VAMP1, 3, 7, 8 and myc-tagged Sec22b and Ykt6 were expressed in HEK293 cells via transient transfection. Cell lysates were incubated with X-LC (FIG. 2G). Both VAMP1 and 3 were cleaved by X-LC, as evidenced by the shift of immunoblot signal to lower molecular weight, while VAMP7, VAMP8, and Sec22B were resistant to X-LC (FIG. 2G).

Unexpectedly, Ykt6 was cleaved by X-LC (FIG. 2G). This finding was confirmed using purified GST-tagged Ykt6 fragment, which shifted to a lower molecular weight band after incubation with X-LC (FIG. 2H). The cleavage site was determined to be K173-S174 by mass spectrometry analysis of the intact Ykt6 versus the Ykt6 cleaved by X-LC (FIG. 13A). This is the homologous site to the cleavage site in VAMP2 (FIG. 2F), indicating that the location of the cleavage site is conserved across different VAMPs. Among VAMP members, VAMP4 contains the same pair of residues (K87-S88) at this site as Ykt6. It was found that GST-tagged cytoplasmic domain of VAMP4 was efficiently cleaved by X-LC (FIG. 21). Consistently, X-LC cleaved native VAMP4 in BDE (FIG. 4J). As a control, Sec22b was not cleaved by X-LC in BDE. In addition, GST-tagged cytoplasmic domain of VAMP5 was also cleaved, although at a slower rate than VAMP2 and VAMP4 (FIG. 21). The cleavage sites were confirmed to be K87-S88 in VAMP4 and R40-S41 in VAMP5 by mass spectrometry analysis (FIG. 14). Both are the homologous sites to the cleavage site in VAMP2 (FIG. 2F). The ability of X-LC to cleave VAMP4, VAMP5, and Ykt6 is highly unusual, as their sequences are substantially different from VAMP 1/2/3. BoNT/X is the first BoNT can cleave VAMPs beyond the canonical targets VAMP1/2/3 66. X-LC also cleaved VAMP4 in BDE, and the cleavage was blocked by EDTA (FIG. 2J).

A remarkable feature of BoNT/X is its unique ability to cleave VAMP4 and Ykt6. VAMP4 is widely expressed and is known to mediate vesicle fusion between trans-Golgi network (TGN) and endosomes, as well as homotypic fusion of endosomes 59, 60. Ykt6 is an atypical SNARE without a transmembrane domain 67-70. It is anchored to membranes via lipidation, which allows dynamic regulation of its membrane association. Ykt6 is an essential protein in yeast, implicated in multiple membrane fusion events including ER-Golgi, intra-
Golgi, endosome-Golgi-vacuolar, and autophagosome formation. Its function in mammalian cells remains to be established. BoNTs are traditionally known to be limited to target SNAREs that mediate vesicle exocytosis onto plasma membranes. BoNT/X is the first BoNT that is capable of cleaving SNAREs mediating various intracellular membrane trafficking events. Interestingly, both VAMP4 and Ykt6 are enriched in neurons. Recent studies suggested that VAMP4 may also contribute to asynchronous synaptic vesicle exocytosis, enlargeosome exocytosis, and activity-dependent bulk endocytosis (ADBE) in neurons. The role of Ykt6 in neurons remains to be established, but it has been shown to suppress the toxicity of α-synuclein in Parkinson's disease models. The other substrate of BoNT/X, VAMP5, is mainly expressed in muscle cells and its function remains to be established. BoNT/X will be a powerful tool for investigating VAMP4, Ykt6, and VAMP5 functions and related membrane trafficking events. In addition, VAMP4 has been implicated in granule release in immune cells, thus BoNT/X might have a unique potential among all BoNTs to modulate inflammatory secretion in immune cells.

**Proteolytic activation of BoNT/X**

BoNTs are initially produced as a single polypeptide. The linker region between LC and H₅ needs to be cleaved by either bacterial or host proteases in a process known as "activation", which is essential for the activity of BoNTs. LC and H₅ of BoNTs remain connected via an inter-chain disulfide bond prior to translocation of LC into the cytosol of cells, where the disulfide bond is reduced in order to release the LC into the cytosol. Sequence alignment revealed that BoNT/X contains the longest linker region between two conserved cysteines compared to all other BoNTs (C423-C467, FIG. 3A). In addition, the linker region of BoNT/X contains an additional cysteine (C461), which is unique to BoNT/X.

To examine whether the linker region between the LC and H₅ of BoNT/X is susceptible to proteolytic cleavage, a recombinant X-LC-H₅ fragment (residues 1-891) was produced in *E. coli* and subjected to limited proteolysis by endoproteinase Lys-C, which cuts at the C-terminal side of lysine residues. To identify the susceptible cleavage site under limited proteolysis conditions, X-LC-H₅ was analyzed using Tandem Mass Tag (TMT) labeling and tandem mass spectrometry approach. TMT labels free N-terminus (and lysines). Limited proteolysis by Lys-C produces additional free N-termini, which would not exist in intact X-LC-H₅ sample (see below for details). Briefly, intact X-LC-H₅ samples were labeled with the light TMT and equal amount of X-LC-H₅ samples were exposed to Lys-C and then labeled...
with the heavy TMT. Both samples were then digested with chymotrypsin, combined together, and subjected to quantitative mass spectrometry analysis. A list of identified peptides was shown in Table 2, below. The light TMT: heavy TMT ratios were usually within 2-fold of each other for each peptide, with the exception for 5 peptides starting with N439, which showed no signal for the light TMT labeling, indicating that N439 is a new N-terminal generated by Lys-C cutting (FIG. 3A, Table 2). Thus, Lys-C preferentially cuts K438-N439 under limited proteolytic conditions, demonstrating that the linker region is susceptible to proteases (FIG. 3A).

Whether this proteolytic activation is important for the function of BoNT/X was examined next. It has been previously shown that incubation of high concentrations of LC-H_N of BoNTs with cultured neurons resulted in entry of LC-H_N into neurons, likely through non-specific uptake into neurons \(^{46,47}\). Using this approach, the potency of intact versus activated X-LC-H_N on cultured rat cortical neurons was compared. Neurons were exposed to X-LC-H_N in media for 12 hours. Cell lysates were harvested and immunoblot analysis was carried out to examine cleavage of SNARE proteins. As shown in FIG. 3B, X-LC-H_N entered neurons and cleaved VAMP2 in a concentration-dependent manner. X-LC-H_N activated by Lys-C showed a drastically increased potency than intact X-LC-H_N. 10 nM activated X-LC-H_N cleaved similar levels of VAMP2 as 150 nM intact X-LC-H_N (FIG. 3B). Note that the intact X-LC-H_N is likely susceptible to proteolytic cleavage by cell surface proteases, which is why it is still active on neurons at high concentrations. Interestingly, activated X-LC-H_N appears to be more potent than activated LC-H_N of BoNT/A (A-LC-H_N) and BoNT/B (B-LC-H_N), which did not show any detectable cleavage of their SNARE substrates in neurons under the same assay conditions (FIG. 3B).

Table 2. Peptide fragments ofX-LC-H\(^{N}\) under limited proteolysis analyzed by TMT labeling and quantitative mass spectrometry.

His6-tagged recombinant X-LC-H\(^{N}\) was labeled with the light TMT. Equal amount of X-LC-H\(^{N}\) samples were exposed to Lys-C and then labeled with the heavy TMT. Both samples were then digested with chymotrypsin, combined together, and subjected to quantitative mass spectrometry analysis. A list of identified peptides was shown. The light TMT: heavy TMT ratios are within 2-fold of each other for all peptide, except five peptides (underlined) starting with N439. These five peptides showed no signal for the light TMT labeling, indicating that N439 is a new N-terminal generated by Lys-C cutting. The peptide sequences in Table 2
correspond, from top to bottom, to SEQ ID NOs: 94-226.

<table>
<thead>
<tr>
<th>ScanF</th>
<th>z</th>
<th>Theo m/z</th>
<th>PPM</th>
<th>X Corr</th>
<th>Δ Corr</th>
<th>Ref.</th>
<th>Peptide</th>
<th>Start Pos</th>
<th>End Pos</th>
<th>Max Heavy</th>
<th>Max Light</th>
<th>H/L ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>16633</td>
<td>2</td>
<td>611.384</td>
<td>1.46</td>
<td>1.751</td>
<td>0.651</td>
<td>xLeH</td>
<td>K.L[IE]NK#F.N</td>
<td>3</td>
<td>8</td>
<td>4.03E+04</td>
<td>2.92E+04</td>
<td>1.38</td>
</tr>
<tr>
<td>16638</td>
<td>2</td>
<td>606.3736</td>
<td>0.95</td>
<td>1.71</td>
<td>0.654</td>
<td>xLeH</td>
<td>K.LE[IN]KF#N</td>
<td>3</td>
<td>8</td>
<td>4.03E+04</td>
<td>2.92E+04</td>
<td>1.38</td>
</tr>
<tr>
<td>15946</td>
<td>2</td>
<td>912.456</td>
<td>2.74</td>
<td>4.083</td>
<td>0.877</td>
<td>xLeH</td>
<td>F.N[IV]NP#IDGINVT M^* R</td>
<td>9</td>
<td>22</td>
<td>4.34E+05</td>
<td>9.26E+05</td>
<td>0.47</td>
</tr>
<tr>
<td>15942</td>
<td>2</td>
<td>909.9508</td>
<td>2.28</td>
<td>4.546</td>
<td>0.661</td>
<td>xLeH</td>
<td>F.NY[NDP]IDGINVT M^* R</td>
<td>9</td>
<td>22</td>
<td>4.34E+05</td>
<td>9.26E+05</td>
<td>0.47</td>
</tr>
<tr>
<td>17201</td>
<td>2</td>
<td>785.9092</td>
<td>1.26</td>
<td>2.455</td>
<td>0.273</td>
<td>xLeH</td>
<td>F.NY[NDDP]IDGINVT</td>
<td>9</td>
<td>20</td>
<td>1.14E+06</td>
<td>1.94E+06</td>
<td>0.59</td>
</tr>
<tr>
<td>11082</td>
<td>2</td>
<td>679.833</td>
<td>1.17</td>
<td>1.74</td>
<td>0.754</td>
<td>xLeH</td>
<td>F.NY[NDDP]IDGIN.V</td>
<td>9</td>
<td>18</td>
<td>8.01E+05</td>
<td>1.20E+06</td>
<td>0.67</td>
</tr>
<tr>
<td>11083</td>
<td>2</td>
<td>682.3382</td>
<td>0.96</td>
<td>2.6</td>
<td>0.742</td>
<td>xLeH</td>
<td>F.N[IV]NP#IDGIN.V</td>
<td>9</td>
<td>18</td>
<td>8.01E+05</td>
<td>1.20E+06</td>
<td>0.67</td>
</tr>
<tr>
<td>19628</td>
<td>2</td>
<td>535.3264</td>
<td>0.97</td>
<td>1.391</td>
<td>0.701</td>
<td>xLeH</td>
<td>D.P[ID]GIN VLT</td>
<td>13</td>
<td>20</td>
<td>6.55E+04</td>
<td>1.23E+05</td>
<td>0.53</td>
</tr>
<tr>
<td>19626</td>
<td>2</td>
<td>532.8211</td>
<td>0.68</td>
<td>1.474</td>
<td>0.608</td>
<td>xLeH</td>
<td>D.P[ID]GIN VLT</td>
<td>13</td>
<td>20</td>
<td>6.55E+04</td>
<td>1.23E+05</td>
<td>0.53</td>
</tr>
<tr>
<td>20815</td>
<td>4</td>
<td>802.2073</td>
<td>2.19</td>
<td>2.962</td>
<td>0.582</td>
<td>xLeH</td>
<td>Y.N[PNYL]NT[PSE]K#DEFLQGV IK#VLE</td>
<td>78</td>
<td>99</td>
<td>4.57E+04</td>
<td>1.21E+05</td>
<td>0.38</td>
</tr>
<tr>
<td>20463</td>
<td>4</td>
<td>802.2073</td>
<td>1.96</td>
<td>2.726</td>
<td>0.385</td>
<td>xLeH</td>
<td>Y.N[PNYL]NT[PSE]K#DEFLQGV IK#VLE</td>
<td>78</td>
<td>99</td>
<td>2.96E+04</td>
<td>8.18E+04</td>
<td>0.36</td>
</tr>
<tr>
<td>20799</td>
<td>4</td>
<td>798.4495</td>
<td>1.91</td>
<td>3.647</td>
<td>0.659</td>
<td>xLeH</td>
<td>Y.NP[NYL]NT[PSE]KD EFLQGVI KVLE</td>
<td>78</td>
<td>99</td>
<td>4.57E+04</td>
<td>1.21E+05</td>
<td>0.38</td>
</tr>
<tr>
<td>20568</td>
<td>4</td>
<td>798.4495</td>
<td>1.63</td>
<td>2.639</td>
<td>0.419</td>
<td>xLeH</td>
<td>Y.NP[NYL]NT[PSE]KD</td>
<td>78</td>
<td>99</td>
<td>2.96E+04</td>
<td>8.18E+04</td>
<td>0.36</td>
</tr>
<tr>
<td>N</td>
<td>EPLQGVI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>KVL.E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22720</td>
<td>2</td>
<td>753.4631</td>
<td>1.96</td>
<td>2.339</td>
<td>0.222</td>
<td>xLeH</td>
<td>N</td>
<td>L.ELISSSI PLPL.V</td>
<td>112</td>
<td>123</td>
<td>1.40E+04</td>
<td>2.80E+04</td>
</tr>
<tr>
<td>21170</td>
<td>2</td>
<td>696.9211</td>
<td>1.9</td>
<td>1.781</td>
<td>0.326</td>
<td>xLeH</td>
<td>N</td>
<td>L.ELISSSI PLPL.V</td>
<td>113</td>
<td>123</td>
<td>2.75E+04</td>
<td>4.15E+04</td>
</tr>
<tr>
<td>21281</td>
<td>2</td>
<td>696.9211</td>
<td>1.86</td>
<td>2.099</td>
<td>0.282</td>
<td>xLeH</td>
<td>N</td>
<td>L.ELISSSI PLPL.V</td>
<td>113</td>
<td>123</td>
<td>2.75E+04</td>
<td>4.15E+04</td>
</tr>
<tr>
<td>21378</td>
<td>2</td>
<td>696.9211</td>
<td>1.83</td>
<td>1.593</td>
<td>0.149</td>
<td>xLeH</td>
<td>N</td>
<td>L.ELISSSI PLPL.V</td>
<td>113</td>
<td>123</td>
<td>2.75E+04</td>
<td>4.15E+04</td>
</tr>
<tr>
<td>19246</td>
<td>2</td>
<td>578.363</td>
<td>1.18</td>
<td>1.443</td>
<td>0.27</td>
<td>xLeH</td>
<td>N</td>
<td>L.ISSIPL PL.V</td>
<td>115</td>
<td>123</td>
<td>1.61E+04</td>
<td>4.12E+04</td>
</tr>
<tr>
<td>19365</td>
<td>2</td>
<td>578.363</td>
<td>1.08</td>
<td>1.624</td>
<td>0.135</td>
<td>xLeH</td>
<td>N</td>
<td>L.ISSIPL PL.V</td>
<td>115</td>
<td>123</td>
<td>1.61E+04</td>
<td>4.12E+04</td>
</tr>
<tr>
<td>19241</td>
<td>2</td>
<td>575.8577</td>
<td>1</td>
<td>1.484</td>
<td>0.209</td>
<td>xLeH</td>
<td>N</td>
<td>L.ISSIPL PL.V</td>
<td>115</td>
<td>123</td>
<td>1.61E+04</td>
<td>4.12E+04</td>
</tr>
<tr>
<td>19360</td>
<td>2</td>
<td>575.8577</td>
<td>0.91</td>
<td>1.673</td>
<td>0.298</td>
<td>xLeH</td>
<td>N</td>
<td>L.ISSIPL PL.V</td>
<td>115</td>
<td>123</td>
<td>1.61E+04</td>
<td>4.12E+04</td>
</tr>
<tr>
<td>13952</td>
<td>2</td>
<td>948.9912</td>
<td>2.47</td>
<td>2.562</td>
<td>0.54</td>
<td>xLeH</td>
<td>N</td>
<td>L.VSNGA LTLSDNEN TIAY.Q</td>
<td>124</td>
<td>139</td>
<td>2.51E+05</td>
<td>3.67E+05</td>
</tr>
<tr>
<td>13949</td>
<td>2</td>
<td>946.486</td>
<td>2.42</td>
<td>1.729</td>
<td>0.599</td>
<td>xLeH</td>
<td>N</td>
<td>L.VSNGA LTLSDNEN TIAY.Q</td>
<td>124</td>
<td>139</td>
<td>2.51E+05</td>
<td>3.67E+05</td>
</tr>
<tr>
<td>6712</td>
<td>2</td>
<td>392.7318</td>
<td>0.19</td>
<td>1.523</td>
<td>0.206</td>
<td>xLeH</td>
<td>N</td>
<td>L.VSNGA LT</td>
<td>124</td>
<td>129</td>
<td>1.18E+05</td>
<td>2.03E+05</td>
</tr>
<tr>
<td>10243</td>
<td>2</td>
<td>678.3482</td>
<td>0.73</td>
<td>1.499</td>
<td>0.616</td>
<td>xLeH</td>
<td>N</td>
<td>L.TLSDNE ETIAY.Q</td>
<td>130</td>
<td>139</td>
<td>2.14E+05</td>
<td>3.67E+05</td>
</tr>
<tr>
<td>10242</td>
<td>2</td>
<td>675.843</td>
<td>0.69</td>
<td>1.594</td>
<td>0.842</td>
<td>xLeH</td>
<td>N</td>
<td>L.TLSDNEN TIAY.Q</td>
<td>130</td>
<td>139</td>
<td>2.14E+05</td>
<td>3.67E+05</td>
</tr>
<tr>
<td>15890</td>
<td>2</td>
<td>1110.579</td>
<td>2.74</td>
<td>2.243</td>
<td>0.6</td>
<td>xLeH</td>
<td>N</td>
<td>L.QJANLVI YGPPDPDI ANNATY. G</td>
<td>150</td>
<td>168</td>
<td>5.20E+04</td>
<td>9.82E+04</td>
</tr>
<tr>
<td>15881</td>
<td>2</td>
<td>1108.073</td>
<td>1.86</td>
<td>2.121</td>
<td>0.673</td>
<td>xLeH</td>
<td>N</td>
<td>L.QANLVI YGPPDPDI ANNATY.</td>
<td>150</td>
<td>168</td>
<td>5.20E+04</td>
<td>9.82E+04</td>
</tr>
<tr>
<td>ID</td>
<td>Value</td>
<td>Weight</td>
<td>MW</td>
<td>PDI</td>
<td>HIC</td>
<td>ID</td>
<td>Value</td>
<td>Weight</td>
<td>MW</td>
<td>PDI</td>
<td>HIC</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>11142</td>
<td>2</td>
<td>727.38</td>
<td>1.39</td>
<td>1.945</td>
<td>0.635</td>
<td>xLcH</td>
<td>L.VIYGPG</td>
<td>PDIANN.A</td>
<td>154</td>
<td>165</td>
<td>1.91E+04</td>
<td>3.90E+04</td>
</tr>
<tr>
<td>12879</td>
<td>2</td>
<td>894.96</td>
<td>1.24</td>
<td>2.673</td>
<td>0.731</td>
<td>xLcH</td>
<td>L.VIYGPG</td>
<td>PDIANNA TY.G</td>
<td>154</td>
<td>168</td>
<td>6.62E+05</td>
<td>1.25E+06</td>
</tr>
<tr>
<td>10964</td>
<td>2</td>
<td>707.35</td>
<td>2.8</td>
<td>1.39</td>
<td>0.466</td>
<td>xLcH</td>
<td>Y.GPGDI ANNATY. G</td>
<td>157</td>
<td>168</td>
<td>1.42E+04</td>
<td>2.24E+04</td>
<td>0.63</td>
</tr>
<tr>
<td>11091</td>
<td>2</td>
<td>456.24</td>
<td>0.78</td>
<td>1.352</td>
<td>0.765</td>
<td>xLcH</td>
<td>N.ATYGL Y.S</td>
<td>166</td>
<td>171</td>
<td>3.55E+04</td>
<td>8.08E+04</td>
<td>0.44</td>
</tr>
<tr>
<td>17435</td>
<td>2</td>
<td>1094.0</td>
<td>2.53</td>
<td>3.418</td>
<td>0.69</td>
<td>xLcH</td>
<td>Y.GJLYST PISNPEGMT LSEVSF.S</td>
<td>169</td>
<td>187</td>
<td>1.55E+05</td>
<td>2.27E+05</td>
<td>0.68</td>
</tr>
<tr>
<td>17410</td>
<td>2</td>
<td>1091.5</td>
<td>2.48</td>
<td>3.748</td>
<td>0.738</td>
<td>xLcH</td>
<td>Y.GLYSTP ISNPEGGT LSEVSF.S</td>
<td>169</td>
<td>187</td>
<td>1.55E+05</td>
<td>2.27E+05</td>
<td>0.68</td>
</tr>
<tr>
<td>19830</td>
<td>2</td>
<td>1259.6</td>
<td>2.15</td>
<td>2.885</td>
<td>0.676</td>
<td>xLcH</td>
<td>Y.GJLYST PISNPEGGT LSEVSFSFSP F.Y</td>
<td>169</td>
<td>190</td>
<td>1.98E+04</td>
<td>2.46E+04</td>
<td>0.81</td>
</tr>
<tr>
<td>20131</td>
<td>3</td>
<td>838.41</td>
<td>2.08</td>
<td>2.781</td>
<td>0.742</td>
<td>xLcH</td>
<td>Y.GLYSTP ISNPEGGT LSEVSFSFSP F.Y</td>
<td>169</td>
<td>190</td>
<td>4.60E+05</td>
<td>5.32E+05</td>
<td>0.86</td>
</tr>
<tr>
<td>12546</td>
<td>2</td>
<td>862.94</td>
<td>1.9</td>
<td>1.817</td>
<td>0.401</td>
<td>xLcH</td>
<td>Y.GJLYST PISNPEGGT LSE.VESL.E</td>
<td>169</td>
<td>183</td>
<td>7.95E+04</td>
<td>7.83E+04</td>
<td>1.02</td>
</tr>
<tr>
<td>12571</td>
<td>2</td>
<td>860.44</td>
<td>1.84</td>
<td>1.528</td>
<td>0.588</td>
<td>xLcH</td>
<td>Y.GLYSTP ISNPEGGT LSE.E</td>
<td>169</td>
<td>183</td>
<td>7.95E+04</td>
<td>7.83E+04</td>
<td>1.02</td>
</tr>
<tr>
<td>19819</td>
<td>3</td>
<td>838.41</td>
<td>1.61</td>
<td>2.655</td>
<td>0.662</td>
<td>xLcH</td>
<td>Y.GLYSTP ISNPEGGT LSEVSFSFSP F.Y</td>
<td>169</td>
<td>190</td>
<td>4.13E+04</td>
<td>4.59E+04</td>
<td>0.90</td>
</tr>
<tr>
<td>9948</td>
<td>2</td>
<td>851.43</td>
<td>2.54</td>
<td>2.044</td>
<td>0.765</td>
<td>xLcH</td>
<td>Y.STPISN GEGTLSE VS.F</td>
<td>172</td>
<td>186</td>
<td>1.92E+04</td>
<td>2.39E+04</td>
<td>0.81</td>
</tr>
<tr>
<td>18409</td>
<td>2</td>
<td>1090.5</td>
<td>2.45</td>
<td>4.434</td>
<td>0.844</td>
<td>xLcH</td>
<td>Y.STPISN GEGTLSE VSFSF.P.Y</td>
<td>172</td>
<td>190</td>
<td>8.21E+05</td>
<td>9.22E+05</td>
<td>0.89</td>
</tr>
<tr>
<td>18417</td>
<td>2</td>
<td>1090.5</td>
<td>2.39</td>
<td>3.307</td>
<td>0.82</td>
<td>xLcH</td>
<td>Y.STPISN GEGTLSE VSFSF.P.Y</td>
<td>172</td>
<td>190</td>
<td>8.21E+05</td>
<td>9.22E+05</td>
<td>0.89</td>
</tr>
<tr>
<td>18418</td>
<td>2</td>
<td>1093.0</td>
<td>2.39</td>
<td>2.983</td>
<td>0.845</td>
<td>xLcH</td>
<td>Y.S]TPISN GEGTLSE</td>
<td>172</td>
<td>190</td>
<td>8.21E+05</td>
<td>9.22E+05</td>
<td>0.89</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>------</td>
<td>-----------------------------</td>
<td>------</td>
<td>------------------------</td>
<td>------</td>
<td>------------------------</td>
<td>------</td>
<td>------------------------</td>
<td>------</td>
<td>------------------------</td>
<td>------</td>
</tr>
<tr>
<td>18876</td>
<td>20193.047</td>
<td>1.97</td>
<td>1.841 0.679</td>
<td>172</td>
<td>190 8.21E+05 9.22E+05 0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14591</td>
<td>2027.4701</td>
<td>1.96</td>
<td>3.362 0.646</td>
<td>172</td>
<td>187 2.09E+05 1.42E+05 1.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14803</td>
<td>20924.9649</td>
<td>1.88</td>
<td>1.934 0.462</td>
<td>172</td>
<td>187 2.96E+05 4.65E+05 0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14852</td>
<td>20924.9649</td>
<td>1.87</td>
<td>1.538 0.369</td>
<td>172</td>
<td>187 2.96E+05 4.65E+05 0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18554</td>
<td>20103.047</td>
<td>1.72</td>
<td>3.044 0.86</td>
<td>172</td>
<td>190 8.21E+05 9.22E+05 0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18552</td>
<td>20109.541</td>
<td>1.68</td>
<td>3.643 0.813</td>
<td>172</td>
<td>190 8.21E+05 9.22E+05 0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14978</td>
<td>20927.4701</td>
<td>1.68</td>
<td>1.401 0.31</td>
<td>172</td>
<td>187 4.96E+04 7.37E+04 0.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18680</td>
<td>20103.047</td>
<td>1.66</td>
<td>1.656 0.717</td>
<td>172</td>
<td>190 8.21E+05 9.22E+05 0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20660</td>
<td>20109.541</td>
<td>1.63</td>
<td>2.182 0.831</td>
<td>172</td>
<td>190 5.24E+05 5.44E+05 0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17226</td>
<td>20924.9649</td>
<td>1.62</td>
<td>2.459 0.698</td>
<td>172</td>
<td>187 3.25E+05 5.31E+05 0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17306</td>
<td>20924.9649</td>
<td>1.62</td>
<td>2.126 0.588</td>
<td>172</td>
<td>187 3.25E+05 5.31E+05 0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10139</td>
<td>20696.3644</td>
<td>1.61</td>
<td>1.942 0.628</td>
<td>172</td>
<td>183 6.39E+04 8.46E+04 0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18674</td>
<td>20109.541</td>
<td>1.59</td>
<td>2.423 0.765</td>
<td>172</td>
<td>190 8.21E+05 9.22E+05 0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20642</td>
<td>20109.541</td>
<td>1.55</td>
<td>2.608 0.823</td>
<td>172</td>
<td>190 5.24E+05 5.44E+05 0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8077</td>
<td>20696.3644</td>
<td>1.54</td>
<td>1.763 0.707</td>
<td>172</td>
<td>183 1.09E+05 1.29E+05 0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18476</td>
<td>20562.2872</td>
<td>1.17</td>
<td>1.812 0.784</td>
<td>183</td>
<td>190 4.49E+04 7.11E+04 0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Mz</td>
<td>Charge</td>
<td>RawScore</td>
<td>Match</td>
<td>Prec</td>
<td>Retention Time</td>
<td>Confidence</td>
<td>Probability</td>
<td>Score</td>
<td>Sequence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>----------</td>
<td>-------</td>
<td>------</td>
<td>----------------</td>
<td>------------</td>
<td>-------------</td>
<td>--------</td>
<td>---------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18478</td>
<td>564.7924</td>
<td>0.95</td>
<td>1.57</td>
<td>0.965</td>
<td>183</td>
<td>4.49E+04</td>
<td>7.11E+04</td>
<td>0.63</td>
<td></td>
<td>L.SJEVSFSPF.Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13190</td>
<td>652.0048</td>
<td>0.97</td>
<td>2.099</td>
<td>0.697</td>
<td>191</td>
<td>1.98E+06</td>
<td>3.09E+06</td>
<td>0.64</td>
<td></td>
<td>F.Y</td>
<td>LK#PF DESYGNYR</td>
<td></td>
</tr>
<tr>
<td>12839</td>
<td>540.6289</td>
<td>0.38</td>
<td>2.441</td>
<td>0.423</td>
<td>191</td>
<td>2.43E+05</td>
<td>5.45E+05</td>
<td>0.44</td>
<td></td>
<td>F.Y</td>
<td>LK#PF DESY.G</td>
<td></td>
</tr>
<tr>
<td>7115</td>
<td>538.7403</td>
<td>0.69</td>
<td>1.476</td>
<td>0.615</td>
<td>196</td>
<td>1.44E+05</td>
<td>3.14E+05</td>
<td>0.46</td>
<td></td>
<td>F.D</td>
<td>JESYGNYR</td>
<td></td>
</tr>
<tr>
<td>16124</td>
<td>661.3911</td>
<td>1.14</td>
<td>2.123</td>
<td>0.575</td>
<td>200</td>
<td>2.81E+04</td>
<td>2.27E+04</td>
<td>1.24</td>
<td></td>
<td>Y.G</td>
<td>NYRS LNVNVRK#F.Y</td>
<td></td>
</tr>
<tr>
<td>11646</td>
<td>474.5894</td>
<td>0.49</td>
<td>2.077</td>
<td>0.602</td>
<td>235</td>
<td>2.80E+05</td>
<td>4.38E+05</td>
<td>0.64</td>
<td></td>
<td>H.NLYGISNRNF.Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10091</td>
<td>603.8103</td>
<td>1.23</td>
<td>2.98</td>
<td>0.756</td>
<td>245</td>
<td>3.05E+05</td>
<td>1.94E+05</td>
<td>1.57</td>
<td></td>
<td>F.Y</td>
<td>YNFD GIKIETSRRQQN,S</td>
<td></td>
</tr>
<tr>
<td>9932</td>
<td>603.8103</td>
<td>1.08</td>
<td>2.5</td>
<td>0.638</td>
<td>245</td>
<td>3.05E+05</td>
<td>1.94E+05</td>
<td>1.57</td>
<td></td>
<td>F.Y</td>
<td>YNFD GIKIETSRRQQN.S</td>
<td></td>
</tr>
<tr>
<td>10782</td>
<td>820.4359</td>
<td>0.65</td>
<td>2.156</td>
<td>0.438</td>
<td>246</td>
<td>7.40E+04</td>
<td>1.33E+05</td>
<td>0.56</td>
<td></td>
<td>Y.Y</td>
<td>NFDTGK#IETSRRQQNLS1,I</td>
<td></td>
</tr>
<tr>
<td>15039</td>
<td>819.1317</td>
<td>1.41</td>
<td>2.281</td>
<td>0.516</td>
<td>298</td>
<td>1.45E+05</td>
<td>2.76E+05</td>
<td>0.53</td>
<td></td>
<td>L.ISERLNTVTVENDLLKY.I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13824</td>
<td>619.6838</td>
<td>1.27</td>
<td>2.021</td>
<td>0.44</td>
<td>303</td>
<td>3.10E+04</td>
<td>5.07E+04</td>
<td>0.61</td>
<td></td>
<td>L.NTVTVENDLLKY.I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10975</td>
<td>760.4251</td>
<td>1.45</td>
<td>2.089</td>
<td>0.663</td>
<td>345</td>
<td>4.63E+05</td>
<td>6.60E+05</td>
<td>0.70</td>
<td></td>
<td>F.V</td>
<td>LINESNLAQRF,S</td>
<td></td>
</tr>
<tr>
<td>17696</td>
<td>966.86</td>
<td>2.14</td>
<td>2.704</td>
<td>0.532</td>
<td>363</td>
<td>8.61E+04</td>
<td>1.42E+05</td>
<td>0.61</td>
<td></td>
<td>H.Y</td>
<td>L&amp;K#E RPIDPIYVNILDDNSYS</td>
<td></td>
</tr>
<tr>
<td>20823</td>
<td>827.9299</td>
<td>2.21</td>
<td>2.495</td>
<td>0.806</td>
<td>371</td>
<td>7.80E+04</td>
<td>8.16E+04</td>
<td>0.96</td>
<td></td>
<td>D.P</td>
<td>IVVNILDNDNSYS</td>
<td></td>
</tr>
<tr>
<td>20825</td>
<td>825.4247</td>
<td>2.17</td>
<td>2.958</td>
<td>0.881</td>
<td>371</td>
<td>7.80E+04</td>
<td>8.16E+04</td>
<td>0.96</td>
<td></td>
<td>D.P</td>
<td>IVVNILDNDNSYS</td>
<td></td>
</tr>
<tr>
<td>12614</td>
<td>2</td>
<td>685.3561</td>
<td>1.34</td>
<td>1.843</td>
<td>0.609</td>
<td>xLeH</td>
<td>N</td>
<td>N.1LDDN SYSTL.E</td>
<td>376</td>
<td>385</td>
<td>1.01E+05</td>
<td>2.57E+05</td>
</tr>
<tr>
<td>11272</td>
<td>2</td>
<td>441.7421</td>
<td>1.48</td>
<td>1.795</td>
<td>0.903</td>
<td>xLeH</td>
<td>N</td>
<td>Y.SJTEGR F.N</td>
<td>383</td>
<td>388</td>
<td>4.45E+04</td>
<td>1.36E+05</td>
</tr>
<tr>
<td>11261</td>
<td>2</td>
<td>439.2369</td>
<td>1.08</td>
<td>1.536</td>
<td>0.514</td>
<td>xLeH</td>
<td>N</td>
<td>Y.STLEGF N</td>
<td>383</td>
<td>388</td>
<td>4.45E+04</td>
<td>1.36E+05</td>
</tr>
<tr>
<td>14260</td>
<td>2</td>
<td>441.7421</td>
<td>0.49</td>
<td>1.462</td>
<td>0.848</td>
<td>xLeH</td>
<td>N</td>
<td>Y.SJTEGR F.N</td>
<td>383</td>
<td>388</td>
<td>5.41E+04</td>
<td>1.32E+05</td>
</tr>
<tr>
<td>14246</td>
<td>2</td>
<td>439.2369</td>
<td>0.38</td>
<td>1.46</td>
<td>0.844</td>
<td>xLeH</td>
<td>N</td>
<td>Y.STLEGF N</td>
<td>383</td>
<td>388</td>
<td>5.41E+04</td>
<td>1.32E+05</td>
</tr>
<tr>
<td>13808</td>
<td>2</td>
<td>916.4629</td>
<td>2.48</td>
<td>2.036</td>
<td>0.498</td>
<td>xLeH</td>
<td>N</td>
<td>F.NISSQG SDNPQGQ LLE</td>
<td>389</td>
<td>403</td>
<td>2.86E+04</td>
<td>6.39E+04</td>
</tr>
<tr>
<td>17314</td>
<td>3</td>
<td>815.7238</td>
<td>2.05</td>
<td>2.149</td>
<td>0.621</td>
<td>xLeH</td>
<td>N</td>
<td>F.NISSQG SDNPQGQ LLESSYF.E</td>
<td>389</td>
<td>408</td>
<td>1.26E+05</td>
<td>2.71E+05</td>
</tr>
<tr>
<td>6340</td>
<td>2</td>
<td>803.3788</td>
<td>0.75</td>
<td>2.554</td>
<td>0.721</td>
<td>xLeH</td>
<td>N</td>
<td>F.NISSQG SDNPQGQ L</td>
<td>389</td>
<td>401</td>
<td>1.03E+04</td>
<td>2.21E+04</td>
</tr>
<tr>
<td>11696</td>
<td>2</td>
<td>493.2814</td>
<td>0.67</td>
<td>1.62</td>
<td>0.518</td>
<td>xLeH</td>
<td>N</td>
<td>L.L.JYNAIY R</td>
<td>429</td>
<td>434</td>
<td>1.19E+05</td>
<td>1.87E+05</td>
</tr>
<tr>
<td>11692</td>
<td>2</td>
<td>490.7762</td>
<td>0.6</td>
<td>1.57</td>
<td>0.118</td>
<td>xLeH</td>
<td>N</td>
<td>L.L.JYNAIY R</td>
<td>429</td>
<td>434</td>
<td>1.19E+05</td>
<td>1.87E+05</td>
</tr>
<tr>
<td>14021</td>
<td>3</td>
<td>756.4332</td>
<td>2.14</td>
<td>2.305</td>
<td>0.363</td>
<td>xLeH</td>
<td>N</td>
<td>K.N1YLNN IDLEDK# K#T.T</td>
<td>439</td>
<td>451</td>
<td>6.85E+05</td>
<td>5.12E+03</td>
</tr>
<tr>
<td>12997</td>
<td>3</td>
<td>1009.909</td>
<td>1.93</td>
<td>3.044</td>
<td>0.553</td>
<td>xLeH</td>
<td>N</td>
<td>K.N1YLNN IDLEDK# K#TTSK#T N.V</td>
<td>439</td>
<td>456</td>
<td>4.24E+05</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>12932</td>
<td>3</td>
<td>1009.909</td>
<td>1.92</td>
<td>4.251</td>
<td>0.695</td>
<td>xLeH</td>
<td>N</td>
<td>K.N1YLNN IDLEDK# K#TTSK#T N.V</td>
<td>439</td>
<td>456</td>
<td>4.24E+05</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>14003</td>
<td>3</td>
<td>790.1157</td>
<td>1.59</td>
<td>2.894</td>
<td>0.542</td>
<td>xLeH</td>
<td>N</td>
<td>K.N1YLNN IDLEDK# K#TT.S</td>
<td>439</td>
<td>452</td>
<td>4.40E+06</td>
<td>1.13E+04</td>
</tr>
<tr>
<td>13105</td>
<td>4</td>
<td>729.173</td>
<td>1</td>
<td>3</td>
<td>0.599</td>
<td>xLeH</td>
<td>N</td>
<td>K.N1YLNN IDLEDK# K#TTSK#T</td>
<td>439</td>
<td>455</td>
<td>3.39E+06</td>
<td>3.14E+03</td>
</tr>
<tr>
<td>N.1</td>
<td>11567</td>
<td>3</td>
<td>747.091</td>
<td>1.33</td>
<td>2.711</td>
<td>0.518</td>
<td>NcH</td>
<td>N.YLNNID LEDKKTT S</td>
<td>440</td>
<td>452</td>
<td>1.22E+04</td>
<td>9.52E+04</td>
</tr>
<tr>
<td>N.1</td>
<td>11515</td>
<td>3</td>
<td>747.091</td>
<td>1.33</td>
<td>2.729</td>
<td>0.486</td>
<td>NcH</td>
<td>N.YLNNID LEDKKTT S</td>
<td>440</td>
<td>452</td>
<td>1.22E+04</td>
<td>9.52E+04</td>
</tr>
<tr>
<td>N.1</td>
<td>17001</td>
<td>2</td>
<td>857.9579</td>
<td>1.99</td>
<td>1.967</td>
<td>0.713</td>
<td>NcH</td>
<td>G.CIEVEN KDL.F.L</td>
<td>467</td>
<td>476</td>
<td>3.29E+04</td>
<td>1.64E+05</td>
</tr>
<tr>
<td>N.1</td>
<td>16337</td>
<td>4</td>
<td>941.793</td>
<td>1.76</td>
<td>2.419</td>
<td>0.135</td>
<td>NcH</td>
<td>F.IJSN# DLNDIN LSEEK#IK #PETTVF. F</td>
<td>477</td>
<td>501</td>
<td>1.27E+06</td>
<td>1.06E+06</td>
</tr>
<tr>
<td>N.1</td>
<td>16394</td>
<td>4</td>
<td>941.793</td>
<td>1.72</td>
<td>2.19</td>
<td>0.391</td>
<td>NcH</td>
<td>F.IJSN# DLNDIN LSEEK#IK #PETTVF. F</td>
<td>477</td>
<td>501</td>
<td>1.27E+06</td>
<td>1.06E+06</td>
</tr>
<tr>
<td>N.1</td>
<td>17983</td>
<td>3</td>
<td>993.888</td>
<td>2.4</td>
<td>3.193</td>
<td>0.657</td>
<td>NcH</td>
<td>K.DSLND INLSEEK* IK#PETTV F.F</td>
<td>482</td>
<td>501</td>
<td>3.27E+04</td>
<td>2.59E+03</td>
</tr>
<tr>
<td>N.1</td>
<td>18031</td>
<td>3</td>
<td>955.5456</td>
<td>2.06</td>
<td>3.607</td>
<td>0.764</td>
<td>NcH</td>
<td>D.SJLNDI NLSEEK#I K#PETTV F.F</td>
<td>483</td>
<td>501</td>
<td>4.90E+04</td>
<td>4.12E+04</td>
</tr>
<tr>
<td>N.1</td>
<td>11720</td>
<td>2</td>
<td>834.8912</td>
<td>2.14</td>
<td>1.689</td>
<td>0.726</td>
<td>NcH</td>
<td>L.SNYDFT EANSIPS.I</td>
<td>514</td>
<td>526</td>
<td>2.91E+04</td>
<td>2.50E+04</td>
</tr>
<tr>
<td>N.1</td>
<td>13409</td>
<td>2</td>
<td>1065.513</td>
<td>2.12</td>
<td>4.742</td>
<td>0.769</td>
<td>NcH</td>
<td>L.SJNYDF TEANSPS ISQQ,N</td>
<td>514</td>
<td>530</td>
<td>7.05E+04</td>
<td>7.52E+04</td>
</tr>
<tr>
<td>N.1</td>
<td>8896</td>
<td>2</td>
<td>642.7908</td>
<td>1.42</td>
<td>1.851</td>
<td>0.708</td>
<td>NcH</td>
<td>L.SNYDFT EAN.S</td>
<td>514</td>
<td>522</td>
<td>2.12E+05</td>
<td>2.27E+05</td>
</tr>
<tr>
<td>N.1</td>
<td>8883</td>
<td>2</td>
<td>645.296</td>
<td>1.39</td>
<td>2.362</td>
<td>0.879</td>
<td>NcH</td>
<td>L.SJNYDF TEAN.S</td>
<td>514</td>
<td>522</td>
<td>2.12E+05</td>
<td>2.27E+05</td>
</tr>
<tr>
<td>N.1</td>
<td>17774</td>
<td>2</td>
<td>1051.044</td>
<td>2.7</td>
<td>1.967</td>
<td>0.9</td>
<td>NcH</td>
<td>Y.DFTEA NSIPSIQ QNIL.E</td>
<td>517</td>
<td>533</td>
<td>3.52E+04</td>
<td>3.28E+04</td>
</tr>
<tr>
<td>N.1</td>
<td>17438</td>
<td>2</td>
<td>1051.044</td>
<td>2.55</td>
<td>1.944</td>
<td>0.665</td>
<td>NcH</td>
<td>Y.DFTEA NSIPSIQ QNIL.E</td>
<td>517</td>
<td>533</td>
<td>9.04E+04</td>
<td>8.08E+04</td>
</tr>
<tr>
<td>N.1</td>
<td>12784</td>
<td>2</td>
<td>880.9387</td>
<td>1.87</td>
<td>2.537</td>
<td>0.778</td>
<td>NcH</td>
<td>Y.DFTEA NSIPSIQ</td>
<td>517</td>
<td>530</td>
<td>4.72E+05</td>
<td>5.06E+05</td>
</tr>
<tr>
<td>N</td>
<td>Q.N</td>
<td>12779</td>
<td>16126</td>
<td>12449</td>
<td>12151</td>
<td>20080</td>
<td>11314</td>
<td>11315</td>
<td>11916</td>
<td>7902</td>
<td>7903</td>
<td>14536</td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>2</td>
<td>883.4439</td>
<td>1.73</td>
<td>3.423</td>
<td>0.797</td>
<td>xLeH</td>
<td>N</td>
<td>Y.DFTEA NSIPSISQ Q.N</td>
<td>517</td>
<td>530</td>
<td>4.72E+05</td>
<td>5.06E+05</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>911.7114</td>
<td>1.6</td>
<td>3.934</td>
<td>0.579</td>
<td>xLeH</td>
<td>N</td>
<td>Y.DFTEA NSIPSISQ QN.I</td>
<td>517</td>
<td>545</td>
<td>5.41E+05</td>
<td>2.02E+05</td>
<td>2.67</td>
</tr>
<tr>
<td>2</td>
<td>937.9602</td>
<td>1.56</td>
<td>3.012</td>
<td>0.73</td>
<td>xLeH</td>
<td>N</td>
<td>Y.DFTEA NSIPSISQ QN.I</td>
<td>517</td>
<td>531</td>
<td>4.65E+05</td>
<td>2.02E+05</td>
<td>0.90</td>
</tr>
<tr>
<td>3</td>
<td>1106.923</td>
<td>2.52</td>
<td>3.592</td>
<td>0.677</td>
<td>xLeH</td>
<td>N</td>
<td>N.SJIPSIS QNILERN NEELYEPI RNSLFE</td>
<td>523</td>
<td>548</td>
<td>1.86E+04</td>
<td>1.52E+04</td>
<td>1.23</td>
</tr>
<tr>
<td>3</td>
<td>700.7002</td>
<td>1.59</td>
<td>2.563</td>
<td>0.763</td>
<td>xLeH</td>
<td>N</td>
<td>L.TDSVDE ALSNPNK VY.S</td>
<td>583</td>
<td>597</td>
<td>1.74E+05</td>
<td>1.84E+05</td>
<td>0.95</td>
</tr>
<tr>
<td>3</td>
<td>704.0405</td>
<td>1.56</td>
<td>2.444</td>
<td>0.262</td>
<td>xLeH</td>
<td>N</td>
<td>L.TJDSVDE EALSNPNK K4VY.S</td>
<td>583</td>
<td>597</td>
<td>1.74E+05</td>
<td>1.84E+05</td>
<td>0.95</td>
</tr>
<tr>
<td>3</td>
<td>700.7002</td>
<td>1.52</td>
<td>2.049</td>
<td>0.508</td>
<td>xLeH</td>
<td>N</td>
<td>L.TDSVDE ALSNPNK VY.S</td>
<td>583</td>
<td>597</td>
<td>2.27E+05</td>
<td>2.78E+05</td>
<td>0.82</td>
</tr>
<tr>
<td>2</td>
<td>635.3637</td>
<td>0.25</td>
<td>1.478</td>
<td>0.567</td>
<td>xLeH</td>
<td>N</td>
<td>L.SNPNK VY.S</td>
<td>591</td>
<td>597</td>
<td>6.79E+05</td>
<td>8.90E+05</td>
<td>0.76</td>
</tr>
<tr>
<td>2</td>
<td>640.3742</td>
<td>0.25</td>
<td>1.485</td>
<td>0.282</td>
<td>xLeH</td>
<td>N</td>
<td>L.SJIPNK #VY.S</td>
<td>591</td>
<td>597</td>
<td>6.79E+05</td>
<td>8.90E+05</td>
<td>0.76</td>
</tr>
<tr>
<td>3</td>
<td>927.1457</td>
<td>1.59</td>
<td>2.027</td>
<td>0.553</td>
<td>xLeH</td>
<td>N</td>
<td>Y.SJPFK4# NM#SNTI NSIETGITS FY.I</td>
<td>598</td>
<td>618</td>
<td>1.47E+05</td>
<td>8.07E+04</td>
<td>1.83</td>
</tr>
<tr>
<td>3</td>
<td>923.8054</td>
<td>1.54</td>
<td>2.35</td>
<td>0.708</td>
<td>xLeH</td>
<td>N</td>
<td>Y.SPPKN M#SNTIN SIETGITS TY.I</td>
<td>598</td>
<td>618</td>
<td>1.47E+05</td>
<td>8.07E+04</td>
<td>1.83</td>
</tr>
<tr>
<td>3</td>
<td>813.421</td>
<td>2.03</td>
<td>2.046</td>
<td>0.548</td>
<td>xLeH</td>
<td>N</td>
<td>F.KNM#S NTINSIET GITSTY.I</td>
<td>601</td>
<td>618</td>
<td>2.41E+06</td>
<td>2.42E+06</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>813.421</td>
<td>1.98</td>
<td>2.378</td>
<td>0.482</td>
<td>xLeH</td>
<td>N</td>
<td>F.KNM#S NTINSIET GITSTY.I</td>
<td>601</td>
<td>618</td>
<td>2.41E+06</td>
<td>2.42E+06</td>
<td>1.00</td>
</tr>
<tr>
<td>13079</td>
<td>3</td>
<td>813.421</td>
<td>1.98</td>
<td>2.245</td>
<td>0.617</td>
<td>xLcH</td>
<td>N</td>
<td>F.KNM*</td>
<td>SNTINSIET</td>
<td>GITSTY.I</td>
<td>601</td>
<td>618</td>
</tr>
<tr>
<td>13194</td>
<td>2</td>
<td>915.470</td>
<td>2.12</td>
<td>1.504</td>
<td>0.654</td>
<td>xLcH</td>
<td>N</td>
<td>M.SJ</td>
<td>NTIN</td>
<td>SIETGITS</td>
<td>TY.I</td>
<td>604</td>
</tr>
<tr>
<td>13170</td>
<td>2</td>
<td>912.964</td>
<td>1.99</td>
<td>1.791</td>
<td>0.699</td>
<td>xLcH</td>
<td>N</td>
<td>M.SNTINS</td>
<td>IETGITS</td>
<td>TY.I</td>
<td>604</td>
<td>618</td>
</tr>
<tr>
<td>12405</td>
<td>2</td>
<td>812.427</td>
<td>2.38</td>
<td>1.349</td>
<td>0.781</td>
<td>xLcH</td>
<td>N</td>
<td>N.TINSIET</td>
<td>GITSTY.I</td>
<td>606</td>
<td>618</td>
<td>2.94E+04</td>
</tr>
<tr>
<td>14318</td>
<td>3</td>
<td>837.465</td>
<td>1.53</td>
<td>2.439</td>
<td>0.647</td>
<td>xLcH</td>
<td>N</td>
<td>F.SJ</td>
<td>DETG</td>
<td>K#IDVIDK</td>
<td>#SSDTL.A</td>
<td>632</td>
</tr>
<tr>
<td>19535</td>
<td>2</td>
<td>586.368</td>
<td>0.52</td>
<td>1.753</td>
<td>0.731</td>
<td>xLcH</td>
<td>N</td>
<td>L.A</td>
<td>JVPYTI</td>
<td>GPL.L</td>
<td>649</td>
<td>657</td>
</tr>
<tr>
<td>19564</td>
<td>2</td>
<td>442.268</td>
<td>0.43</td>
<td>1.498</td>
<td>0.892</td>
<td>xLcH</td>
<td>N</td>
<td>V.PYTGPL.L</td>
<td>652</td>
<td>657</td>
<td>1.30E+04</td>
<td>2.64E+04</td>
</tr>
<tr>
<td>9998</td>
<td>2</td>
<td>715.396</td>
<td>1.76</td>
<td>2.148</td>
<td>0.33</td>
<td>xLcH</td>
<td>N</td>
<td>V.I</td>
<td>JGGEL</td>
<td>AREQVE.A</td>
<td>669</td>
<td>709</td>
</tr>
<tr>
<td>5372</td>
<td>3</td>
<td>494.953</td>
<td>0.34</td>
<td>2.943</td>
<td>0.626</td>
<td>xLcH</td>
<td>N</td>
<td>L.SRQAN</td>
<td>AIKM*.N</td>
<td>754</td>
<td>762</td>
<td>2.50E+05</td>
</tr>
<tr>
<td>5363</td>
<td>3</td>
<td>494.953</td>
<td>0.32</td>
<td>3.357</td>
<td>0.734</td>
<td>xLcH</td>
<td>N</td>
<td>L.SRQAN</td>
<td>AIKM*.N</td>
<td>754</td>
<td>762</td>
<td>2.50E+05</td>
</tr>
<tr>
<td>17794</td>
<td>2</td>
<td>734.353</td>
<td>1.89</td>
<td>2.212</td>
<td>0.79</td>
<td>xLcH</td>
<td>N</td>
<td>F.SEFDDL</td>
<td>LNQY.K</td>
<td>879</td>
<td>888</td>
<td>3.00E+05</td>
</tr>
<tr>
<td>14171</td>
<td>4</td>
<td>757.641</td>
<td>0.97</td>
<td>2.289</td>
<td>0.421</td>
<td>xLcH</td>
<td>N</td>
<td>F.DDLINQ</td>
<td>YKNEGSI</td>
<td>LPETGGL.EHH.H</td>
<td>882</td>
<td>904</td>
</tr>
<tr>
<td>7676</td>
<td>4</td>
<td>679.353</td>
<td>1.49</td>
<td>2.624</td>
<td>0.624</td>
<td>xLcH</td>
<td>N</td>
<td>Y.KNEGSI</td>
<td>LPETGGL.EHHHHH.H-</td>
<td>889</td>
<td>908</td>
<td>3.63E+04</td>
</tr>
<tr>
<td>13193</td>
<td>3</td>
<td>588.336</td>
<td>1.18</td>
<td>2.307</td>
<td>0.629</td>
<td>xLcH</td>
<td>N</td>
<td>Y.KNEGSI</td>
<td>LPETGGL.E</td>
<td>889</td>
<td>901</td>
<td>2.81E+05</td>
</tr>
<tr>
<td>10362</td>
<td>3</td>
<td>677.037</td>
<td>1.15</td>
<td>2.582</td>
<td>0.615</td>
<td>xLcH</td>
<td>N</td>
<td>Y.KNEGSI</td>
<td>LPETGGL.EH.H</td>
<td>889</td>
<td>903</td>
<td>8.01E+05</td>
</tr>
<tr>
<td>9325</td>
<td>4</td>
<td>542.294</td>
<td>1.06</td>
<td>2.617</td>
<td>0.58</td>
<td>xLcH</td>
<td>N</td>
<td>Y.KNEGSI</td>
<td>889</td>
<td>904</td>
<td>2.17E+05</td>
<td>2.96E+05</td>
</tr>
</tbody>
</table>
Unique feature of the disulfide bond in BoNT/X

The linker region of BoNT/X contains an additional cysteine (C461), which is unique to BoNT/X. To determine which cysteine forms the disulfide bond connecting the LC and HC, three X-LC-H_N mutants were generated, with each of the three cysteine residues mutated (C423S, C461S, and C467S). These three cysteine mutants, as well as wild type (WT) X-LC-H_N were subjected to limited proteolysis and then analyzed via SDS-PAGE and Coomassie Blue staining, with or without reducing agent DTT (FIG. 3D). It was found that mutating the cysteine on the LC (C423S) resulted in a protein that separated into two ~50 kDa bands, with or without DTT, indicating that C423S abolished the inter-chain disulfide bond. In contrast, mutants containing either C461S or C467S showed a single band at 100 kDa in the absence of DTT, which separated into two -50 kDa bands in the presence of DTT, suggesting that both C461 and C467 on the H_N can form inter-chain disulfide bond with C423 on the LC. Also the X-LC-H_N (C423S) mutant appears to be more susceptible to Lys-C than both C461S and C467S mutants, resulting in further degradation of the protein (FIG. 3C). This result suggests that losing the inter-chain disulfide bond may increase the freedom of the LC and H_N, thus exposing more surface areas. Furthermore, a portion of WT X-LC-H_N formed aggregates at the top of the SDS-PAGE gel (FIG. 3C). These aggregates are due to formation of inter-molecular disulfide bond, as they disappeared in the presence of DTT (FIG. 3C, +DTT). C423, C461 and C467 are the only three cysteines in X-LC-H_N. Mutating any one of these cysteines abolished the X-LC-H_N aggregates (FIG. 3C, -DTT), indicating that formation of inter-molecular disulfide bond is due to existence of an extra cysteine in the linker region.

The majority of activated WT X-LC-H_N also separated to two -50 kDa bands on SDS-PAGE gel without DTT (FIG. 3C). On the other hand, WT X-LC -H_N is similarly resistant to Lys-C as C461S and C467S mutants, and it showed no further degradation as C423S mutant did (FIG. 3C, +DTT), suggesting that WT X-LC -H_N is different from C423S mutant. One possible explanation is disulfide bond shuffling due to the existence of two cysteines close to each on the H_N (C461 and C467), which can rearrange the disulfide bond from inter-chain C423-C467 or C423-C467 to intra-chain C461-C467 under denatured conditions. To test this hypothesis, an alkylating reagent, N-Ethylmaleimide (NEM), which reacts with sulphydryls of free cysteine and permanently block any free cysteines, was used. As shown in FIG. 3D,
WT X-LC-H\textsubscript{N} pretreated with NEM showed largely a single band at 100 kDa in the absence of DTT, and separated into two -50 kDa bands in the presence of DTT. These results confirmed that native WT X-LC-H\textsubscript{N} contains mainly inter-chain disulfide bond, which is susceptible to disulfide bond shuffling due to the existence of the third cysteine in the linker region.

Finally, the activity of the three X-LC-H\textsubscript{N} cysteine mutants on cultured neurons was examined. As shown in FIG. 3E, mutating the cysteine on the LC (C423S) abolished the activity of X-LC-H\textsubscript{N}, as evidenced by lack of VAMP2 cleavage in neurons. Mutating one of the two cysteines on the H\textsubscript{N} (C461 or C467) did not significantly affect the potency of X-LC-H\textsubscript{N} compared to wild type (WT) X-LC-H\textsubscript{N} (FIG. 3E). These results confirmed that the inter-chain disulfide bond is essential for the activity of BoNT/X and demonstrated that functional inter-chain disulfide bond can be formed via either C423-C461 or C423-C467.

**Generating full-length BoNT/X via sortase-mediated ligation**

To evaluate whether BoNT/X is a functional toxin, it was necessary to generate and test full-length BoNT/X. However, BoNTs are one of the most dangerous potential bioterrorism agents. Therefore, the necessary precaution was taken, and the full-length active toxin gene was not generated. Instead, an approach to generate limited amounts of full-length BoNTs in test tubes under controlled conditions by the enzymatic ligation of two non-toxic fragments of BoNTs was developed. This method utilizes a transpeptidase known as sortase, which recognizes specific peptide motifs and covalently link two peptides together by forming a native peptide bond (FIG. 4A). This approach has been previously utilized to generate chimeric toxins and other fusion proteins\textsuperscript{50,51}.

An engineered sortase A, known as SrtA\textsuperscript{*}, from *Staphylococcus aureus* was generated\textsuperscript{51}. SrtA\textsuperscript{*} recognizes the peptide motif LPXTG (SEQ ID NO: 57), cleaves between T-G, and concurrently forms a new peptide bond between the protein containing LPXTG (SEQ ID NO: 57) with other proteins/peptides containing one or more N-terminal glycine (FIG. 4A). Two non-toxic fragments of BoNT/X: (1) LC-H\textsubscript{N} with LPETGG (SEQ ID NO: 58) motif and a His\textsubscript{6}-tag fused to the C-terminus; (2) the He of BoNT/X (X-Hc) with a GST tag and thrombin cleavage site at its N-terminus were produced. Cutting by thrombin releases X-3\textsuperscript{4} with a free glycine at the N-terminus. Incubation of these two fragments with SrtA\textsuperscript{*} generated limited amount of -150 kD full-length BoNT/X containing a short linker (LPETGS, SEQ ID NO: 59) between LC-H\textsubscript{N} and H\textsubscript{c} (FIGs. 4A-4B).
It was observed that X-Hc showed a strong tendency for aggregation in solution for unknown reasons once it is cut from GST tag, which might be the reason why the ligation efficiency is low for BoNT/X (FIG. 4B). In contrast, ligation of X-LC-H₈ with the H₈ of BoNT/A (A-Hc) using the same approach achieved a much higher efficiency, with majority of X-LC-H₈ ligated into a full-length XA chimeric toxin (FIG. 8A).

**BoNT/X is active on cultured neurons**

To analyze the activity of full-length BoNT/X, cultured rat cortical neurons as a model system were used. Neurons were exposed to the sortase ligation mixture and various control mixtures in media. Cell lysates were harvested 12 hours later and immunoblot analysis was carried out to examine cleavage of SNARE proteins. As shown in FIG. 4C, X-LC-H₈ alone cleaved some VAMP2 due to its high concentration in the reaction mixture. The control mixture containing X-LC-H₈ and X-Hc but not sortase slightly enhanced cleavage of VAMP2 compared to X-LC-H₈ alone. This result suggests that X-3/4 might be associated with X-LC-H₈ via non-covalent interactions. This interaction appears to be specific as the control mixture containing X-LC-H₈ and A-Hc showed the same level of VAMP2 cleavage as X-LC-H₈ alone (FIG. 8B). Ligating X-LC-H₈ and X-Hc by sortase enhanced cleavage of VAMP2 over the mixture of X-LC-H₈ and X-Hc without sortase (FIG. 4C), demonstrating that ligated full-length BoNT/X can enter neurons and cleave VAMP2. Similarly, ligated full-length XA chimeric toxin also entered neurons and cleaved VAMP2 (FIG. 8B).

Mixing X-Hc with X-LC-H₈ increased the amounts of aggregates at the top of the SDS-PAGE gel compared to X-LC-H₈ alone. These aggregates disappeared in the presence of DTT, suggesting that a portion of X-Hc formed inter-molecular disulfide bond with X-LC-H₈. The presence of DTT also increased the amount of ligated full-length BoNT/X, suggesting that a portion of BoNT/X aggregated via inter-molecular disulfide bond (FIG. 4B). The formation of these aggregates could significantly reduce the effective toxin monomer concentrations in solution. This could be an intrinsic weakness of BoNT/X sequence. X-Hc contains a single cysteine (C1240) and mutating this cysteine did not affect the activity of ligated BoNT/X (FIG. 9). Furthermore, C1240S mutant can be combined with C461S or C467S mutations in the X-LC-H₈ to generate a modified BoNT/X with no free cysteines (FIG. 9). These mutant toxins maintained the same levels of activity as WT BoNT/X, but are more stable in solution as monomers than WT BoNT/X.
BoNT/X induced flaccid paralysis in vivo in mice

Whether BoNT/X is active in vivo was examined using a well-established non-lethal assay in mice, known as Digit Abduction Score (DAS). This assay measures local muscle paralysis following injection of BoNTs into mouse hind limb muscles. BoNTs cause flaccid paralysis of limb muscles, which can be detected by the failure to spread the toes during the startle response. An activated sortase reaction mixture (FIG. 4B, lane 7) was injected into the gastrocnemius muscles of the right hind limb in mice. Within 12 hours, the right limb developed typical flaccid paralysis and the toes failed to spread (FIG. 4D). These results confirmed that BoNT/X is capable of causing flaccid paralysis in vivo as other BoNTs.

BoNT/X was not recognized by antisera raised against all known BoNTs

To further confirm that BoNT/X is a serologically unique BoNT, dot blot assays were carried out using antisera raised against known BoNTs, including all seven serotypes as well as one mosaic toxin (BoNT/DC). Four horse antisera were utilized (trivalent anti-BoNT/A, B, and E, anti-BoNT/C, anti-BoNT/DC, and anti-BoNT/F), as well as two goat antisera (anti-BoNT/G and anti-BoNT/D). These antisera were all capable of neutralizing their corresponding target BoNTs and prevented cleavage of SNARE proteins in neurons (FIG. 10), thus validating their specificity and potency. As shown in FIG. 4E, these antisera recognized their corresponding target toxins, yet none of them recognized BoNT/X. Note that the antisera raised against BoNT/DC and BoNT/C cross-react with each other, as they share high degree of similarity in their He. These result established BoNT/X as a new serological type of BoNTs.

Full-length inactive BoNT/X

Finally, whether full-length BoNT/X can be produced as a soluble protein was examined. To ensure the biosafety requirement, mutations in the LC of BoNT/X were introduced that inactivate its toxicity. Mutations at two residues R362A/Y365F in BoNT/A have been shown to inactivate the protease activity of the LC in vitro and abolishes the toxicity of full-length BoNT/A in mice in vivo. These two residues are conserved in all BoNTs including BoNT/X. Therefore, the corresponding mutations were introduced at these two sites (R360A/Y363F in BoNT/X). As shown in FIG. 4F, this full-length inactivated form of BoNT/X (BONT/X<sub>RV</sub>) was produced and purified as a His6-tagged protein in E.coli recombinantly. It does not have any activity on neurons as VAMP2 was not cleaved in neurons (FIG. 11).
A substantial portion of BoNT/X<sub>RY</sub> formed aggregates at the top of the SDS-PAGE gel (FIG. 4F). This is likely due to formation of inter-molecular disulfide bond from the extra cysteine in the linker region and the cysteine in the He, as adding DTT converted the aggregates to monomeric BoNT/X<sub>RY</sub> (FIG. 4F). Mutating these cysteines does not affect the activity of BoNT/X (FIG. 9), and has the benefit of preventing formation of inter-molecular disulfide bond and aggregations of BoNT/X.

An inactive form of BoNT/X might be utilized as a vehicle to deliver therapeutics into neurons. Inactivation can be achieved by mutations at any one of the following residues or their combinations: R360, Y363, H227, E228, or H231, with the later three residues forming the conserved protease motif.

**Purification of full-length inactive BoNT/X at industrial-scale**

Whether full-length BoNT/X can be purified to a high degree of purity and with a good yield, which will be important for industrial production of BoNT/X (or its derivative) as a therapeutic toxin, was investigated. Several parameters of cell growth and expression were tested, such as temperature, time of induction and IPTG concentrations. The optimal parameters chosen for protein expression were culture of the cells at 37°C until they reached exponential growth, at which stage the temperature was reduced to 18°C and expression induced by addition of ImM IPTG to the media. Cells were then cultured for 16 to 18 hours before harvesting. Presence of BoNT/X was verified by SDS-PAGE and showed a high level of over-expression in the soluble fraction (FIG. 11B).

Several small-scale purification trials were carried out to optimize the production process. Mechanical cell lysis using an Emulsiflex-C3 (Avestin, Mannheim, Germany) was the preferred method for intracellular protein extraction, and appeared more efficient than sonication. Various buffer conditions also had to be assessed for optimal recovery of BoNT/X. A reducing agent was included throughout the purification process and greatly decreased the propensity to unwanted aggregation. Additionally, glycerol was used as an additive during the early stage of the purification process and improved protein stability.

The BoNT/X construct was expressed with a HIS6-tag that could be used for affinity chromatography as a first purification step. For small-scale trials, a 5ml HiTrapFF column (GE Healthcare, Danderyd, Sweden) was used. In order to achieve the highest purity from the initial chromatography, various concentrations of imidazole were tested. BoNT/X eluted from a concentration of 100mM imidazole; however, a major contaminant readily co-purified with
the toxin. This contaminant appeared to non-specifically interact with BoNT/X and was identified by mass spectrometry as an *E. coli* host protein (bifunctional polymyxin resistance protein ArnA). The presence of this contaminant was dramatically reduced with the introduction of a high salt concentration (500 mM NaCl) and by carrying out an additional washing step at 100 mM imidazole during purification. This allowed for elution of a purer BoNT/X fraction at 250 mM imidazole. This later fraction could then be polished by size exclusion chromatography.

Once in place, this protocol was scaled up by expressing up to 12 L of media with the conditions described above. Additionally, a larger affinity chromatography matrix was prepared consisting of 15 ml of Protino® Ni-NTA agarose (Macherey-Nagel, Diiren, Germany) to increase the yield of BoNT/X recovery. The final purification step was performed by size exclusion chromatography using a Superdex200-16/60 column (GE Healthcare, Danderyd, Sweden). Using this method, between 85 and 90% purity was obtained (FIG. 11C). The protein could be concentrated (using a Vivaspin concentrator with a 100kDa cut-off; GE Healthcare, Danderyd, Sweden) and appeared stable up to 10mg/ml. Complete details of the protein production process are described below. The yield of BoNT/X obtained was approximately 3 mg per liter of cell culture. Together these results demonstrated that BoNT/X can be purified at industrial scale to high purity.

Note that the purification was done in the presence of reducing agent, which would reduce the disulfide bond between the LC and the HC, so purified toxin would not be active. However, a designed BoNT/X derivative containing mutations at the cysteine sites (one mutation at C461 or C467, combined with mutating C1240) would be able to be purified without reducing agents. Note that an inactive form of BoNT/X (and its cysteine mutation derivative) might be utilized as a vehicle to deliver therapeutics into neurons. Inactivation can be achieved by mutations at any one of the following residues or their combinations: R360, Y363, H227, E228, or H231 (the later three residues form the conserved protease motif).

**Identification of gangliosides as receptors for BoNT/X**

Gangliosides are well-established lipid co-receptors for all BoNTs and a ganglioside-binding motif is well-conserved in BoNT/X (FIG. 1C). Highly purified full-length inactive BoNT/X was used to examine whether BoNT/X binds to neuronal cells via gangliosides/ An in vitro ELISA assay was developed to test for interaction with four major brain gangliosides: GDla, GDlb, GTlb, and GM1. A-LC was use as a negative control to assess unspecific
binding. Direct comparison with the receptor binding domain of BoNT/A (A-HC) was also performed. Binding of proteins were detected using an anti-His6-tag antibody. It was found that BoNT/X showed a dose-dependent binding to all four gangliosides over the non-specific binding level of A-LC (FIG. 12), suggesting that BoNT/X is capable of utilizing all four brain gangliosides as co-receptors. In accordance with previous reports, BoNT/A presented an equal preference for GDla and GTlb (FIG 12F) and their terminal NAcGal-Gal-NAcNeu moiety (with apparent EC50 values of 0.7 and 1.0μM, respectively, when fitted with a sigmoidal dose-response model). In contrast, BoNT/X showed higher affinity for GDlb and GM1 over GDla and GTlb (FIG 12E). This would suggest BoNT/X has a preferred sialic acid recognition pattern, also seen in BoNT/B and TeNT. BoNT/X possesses the conserved SxWY motif at a homologous location to the one of the other toxins. The fact that it could recognize all four gangliosides, albeit with low affinity, may be an indication of multiple carbohydrate binding sites.

Discussion

The eighth serotype of BoNTs over 45 years after the identification of the last major BoNT serotype has been identified. BoNT/X has the lowest protein sequence identity to any other BoNTs and TeNT among this family of toxins, and this low level of identity is evenly distributed along the toxin sequence. As expected, BoNT/X was not recognized by any antisera raised against known BoNTs. It clearly represents a unique and distinct evolutionary branch of the toxin family.

BoNT/X was revealed by searching genomic sequences of Clostridium botulinum strains and it represents the first major toxin type identified by genomic sequencing and bioinformatics approach. The strain 111 that contains BoNT/X gene was initially identified from an infant botulism patient in 1990s. Previous characterizations using classic neutralization assay have established BoNT/B2 as the major toxin of this strain. It is likely that BoNT/X is a silent toxin gene, or it was not expressed at detectable toxicity levels under the culture conditions in the lab. Therefore, it can only be identified by sequencing strain 111. This illustrates the importance of genomic sequencing and bioinformatics approaches for understanding microbial virulent factors.

Silent BoNT genes have been frequently found previously in various Clostridium botulinum strains. It is not clear why these bacteria keep a silent toxin gene. It could be an evolutionarily degenerated gene. This is clearly the case when silent toxin genes contain
premature stop code mutations. However, there are also cases that the silent gene encodes a full-length BoNT. Whether these silent full-length BoNTs might be expressed and exhibit toxicity under certain environmental conditions remains an intriguing question.

The general three-domain structures and functions of BoNTs are well conserved in BoNT/X, but it also has a few unique characteristics: (1) it shares VAMP as its target in neurons with BoNT/B, D, F, and G, but it cuts VAMP at a novel site (R66-A67 in VAMP2) that is unique to this toxin. This further expands the repertoire of toxins that can be used to ablate VAMP at different sites. (2) The inter-chain disulfide bond connecting LC and H_N is conserved in BoNT/X, but it also contains a unique additional cysteine in the linker region, which may lead to disulfide bond shuffling. The extra cysteine on H_N is not essential for the activity of LC-H_N (FIG. 3D), and mutating it has the benefit of preventing formation of inter-molecular disulfide bond (FIG. 3D, 4B).

His6-tagged X-LC-H_N fragment are stable in buffers as recombinant proteins. It showed a higher level of activity on neurons than both A-LC-H_N and B-LC-H_N (FIG. 3B), suggesting that its membrane translocation and/or protease activity might be more efficient than the corresponding fragments in BoNT/A and BoNT/B. X-LC-H_N could be a useful reagent for targeting VAMP 1/2/3 in a broad range of cell types and tissues as its entry might not be restricted to neurons. For instance, it potentially can be utilized to reduce pain in a local region by targeting both sensory neurons and other cells that secrete inflammatory signals. It could also be used to generate chimeric toxins, such as XA (FIG. 8).

X-Hc is functional as its presence enhanced cleavage of VAMP2 in neurons than LC-H_N alone (FIG. 4C). However, X-Hc may have some unfavorable characterizes that remain to be further evaluated. For instance, sufficient levels of soluble X-Hc were only produced when it was fused with GST, which is known to facilitate protein folding/solubility, but not with His6 tag. Once released from GST tag, X-Hc is prone to aggregation. In addition, the cysteine in X-Hc may also form inter-molecular disulfide bond (FIG. 4B). Full-length inactive BoNT/X can be purified and exist as a soluble protein, suggesting that the solubility issue with X-Hc might at least partially due to separation of this domain from X-LC-H_N. For instance, X-LC-H_N might interact with X-Hc and covers its potential hydrophobic segments in the full-length context, which is not unusual for a multi-domain protein.

Gangliosides have long been established as neuronal receptors for all BoNT subtypes. It is demonstrated that BoNT/X can bind to all four of the most abundant gangliosides: GDla, GDlb, GTlb, and GM1. Additionally it does so with remarkable difference in affinity and
specificity when compared to BoNT/A. This is an intriguing property, as other BoNTs appear to have various degrees of preferences toward a subgroup of gangliosides. For instance, BoNT/A, E, F, and G prefer GDla and GTlb. BoNT/X might potentially recognize a broader range of neuron types compared to other BoNTs.

It is possible that BoNT/X has a low toxicity in vivo, which might explain why BoNT/X activity was not detected in the original study on strain 111. If this is the case, the reduced toxicity is likely due to its He domain, as X-LC-H requires appears to be more active than both A-LC-H and B-LC-H. The formation of inter-molecular disulfide bond might also reduce the effective toxin concentration. It will be necessary to produce full-length native BoNT/X in order to determine its potency in vivo, but it will be important to generate neutralizing antisera using non-toxic fragments of BoNT/X prior to producing full-length toxin.

Introducing full-length active toxin gene into any expression systems/organisms is always a significant biosafety concern and it has become a formidable hurdle for structure-function studies of biological toxins. This is particularly an important consideration for BoNTs as they are one of the six category A potential bioterrorism agents. Here a method to assemble limited amount of full-length toxin biochemically from two complementary and non-toxic fragments was developed. Each fragment is expressed and purified individually, and then ligated together by sortase in test tubes. Other protein ligation methods such as split intein systems, which fuse two protein fragments through protein trans-splicing, can also be utilized. By controlling the amount of precursor fragments in the reaction, the amount of ligated full-length toxin can be strictly controlled. This "semi-synthesis" approach can be used to produce multi-domain biological toxins and other toxic proteins under controlled conditions. It also provides a versatile platform for generating fusion and chimeric toxins, such as swapping the He of two BoNTs, replacing He of BoNTs with other targeting proteins, or attaching additional cargo to toxins. As there is no full-length toxin cDNA ever generated and no expression of toxins in bacteria or any other living organisms, this approach significantly mitigates the biosafety concerns associated with producing wild type and mutant toxins and will greatly facilitate structure-function studies of biological toxins and toxic proteins.

**Materials and Methods**
**Materials:** Mouse monoclonal antibodies for syntaxin 1 (HPC-1), SNAP-25 (C171.2), and VAMP2 (C169.1) were generously provided by E. Chapman (Madison, WI) and are available from Synaptic Systems (Goettingen, Germany). Mouse monoclonal antibody for actin was purchased from Sigma (AC-15). Equine polyclonal antisera against BoNT/A/B/E, BoNT/C, BoNT/DC, BoNT/F, and goat polyclonal antisera against BoNT/G were obtained from the FDA. Goat polyclonal antibody against BoNT/D was purchased from Fisher Scientific (NB 10062469). BoNT/A, BoNT/B, BoNT/C, BoNT/DC, BoNT/E, BoNT/F, and BoNT/G were purchased from Metabiologics (Madison, WI). BoNT/D was generously provided by E. Johnson (Madison, WI).

**cDNA and constructs:** The cDNAs encoding X-LC (residues 1-439) and X-Hc (residues 893-1306) was synthesized. The cDNA encoding X-HN was generated in-house using Gibson assembly method. The cDNAs encoding A-LC (residues 1-425, M30196) and B-LC (residues 1-439, AB232927) were synthesized by GenScript (New Brunswick, NJ). These LCs were cloned into pET28 vectors for expression as His6-tagged proteins. X-Hc was cloned into pGEX4T to express as a GST-tagged protein. X-LC-H_N, A-LC-H_N, and B-LC-H_N were subcloned into pET28 vector, with a peptide sequence LPETGG (SEQ ID NO: 58) fused to their C-termini, and were purified as His6-tagged proteins. Full-length inactive form of BoNT/X was assembled in-house from mutated X-LC (R360A/Y363F), X-H_N, and X-H_c. It was cloned into pET28 vector with a His6-tagg fused to the C-terminus of BoNT/X. The cDNA encoding rat VAMP2 was generously provided by E. Chapman (Madison, WI). VAMP2 (1-96) was cloned into pET28 vector and expressed as a His6-tagged protein. VAMP2 (33-86) was cloned into a pGEX4T vector and expressed as a GST-tagged protein. The cDNA encoding mouse VAMP1, VAMP3, rat VAMP7, and VAMP8 was generously provided by C. Hu (Louisville, KY). They were cloned into a modified pcDNA3.1 vectors, with a HA tag fused to their C-termini. The construct encoding His6-tagged sortase (SrtA*) was generously provided by B. Pentelute (Boston, MA) and has been described previously.

**Bioinformatics:** The Uniprot database was searched with jackhammer at the HMMER web server using a BoNT type A sequence (Uniprot accession number A5HZZ9) until convergence. Returned sequences were aligned with Clustal Omega and a NeighborNet phylogenetic network estimated with SplitsTree4.

**Protein purification:** *E.coli* BL21 (DE3) was utilized for protein expression. Induction of expression was carried out with 0.1 mM IPTG at 22 °C overnight. Bacterial pellets were disrupted in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl) by sonication and supernatants
were collected after centrifugation at 20000g for 30 min at 4°C. Protein purification was carried out using AKTA Prime FPLC system (GE) and purified proteins were further desalted with PD-10 column (GE, 17-0851-01). Specifically, full-length inactive BoNT/X (BoNT/X<sub>RY</sub>) was cloned into a pET22b vector. The corresponding plasmid was transformed into E. coli BL21 (DE3) competent cells. Resulting colonies were used to inoculate 100 ml overnight cultures of TB medium containing 100 µg/ml Carbenicillin in 250 ml shake-flask and grown at 37°C. Cultures for expression were first grown using a LEX Bioreactor (Epiphyte3, Ontario, Canada) at 37°C in 1.5 L of TB media until OD<sub>600</sub> reached 0.8. The temperature was then reduced to 18°C for induction of expression with 1 mM IPTG, and grown for 16-17 hours.

Cells were harvested and re-suspended on ice in 50 mM HEPES pH 7.2, 500 mM NaCl, 25 mM imidazole, 5% glycerol, 2 mM TCEP to allow for cell lysis with an Emulsiflex-C3 (Avestin, Mannheim, Germany) at 20,000 psi. Lysate was ultra-centrifuged at 200,000 g for 45 minutes at 4°C. Supernatant was loaded onto a 15 ml Protino® Ni-NTA agarose (Macherey-Nagen, Diiren, Germany) column that was then washed with 50 mM HEPES pH 7.2, 500 mM NaCl, 100 mM imidazole, 5% glycerol, 1 mM TCEP. Elution was carried out with 50 mM HEPES pH 7.2, 500 mM NaCl, 250 mM imidazole, 5% glycerol, 1 mM TCEP. The eluate was dialyzed overnight in 50 mM HEPES pH 7.2, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP at 4°C. Dialysate was concentrated using a Vivaspin concentrator (100KD cut-off, GE Healthcare, Danderyd, Sweden) before being loaded on a Superdex200-16/60 column (GE Healthcare, Danderyd, Sweden) pre-equilibrated in the same buffer as was used for dialysis. The elution peak corresponding to BoNT/X was collected and concentrated so that the final sample was at a concentration of 10 mg/ml. The sample was aliquoted and flash-frozen in liquid nitrogen for storage at -80°C.

**Ganglioside binding assay:** Purified gangliosides GDI<sub>a</sub>, GDI<sub>b</sub>, GTI<sub>b</sub>, and GM1 (Carbosynth, Compton, UK) were dissolved in DMSO and diluted in methanol to reach a final concentration of 2.5 µg/ml; 100 µL was applied to each well of a 96-well PVC assay plate (catalog no. 2595, Corning; Corning, NY). After solvent evaporation at 21°C, the wells were washed with 200 µL PBS/0.1% (w/v) BSA. Nonspecific binding sites were blocked by incubation for 2.5 h at 4°C in 200 µL of PBS/2% (w/v) BSA. Binding assays were performed in 100 µL PBS/0.1% (w/v) BSA per well for 1 h at 4°C containing samples (triplicate) at concentrations ranging from 6 µM to 0.05 µM (in serial 2-fold dilution). Following incubation, wells were washed three times with PBS/0.1% (w/v) BSA and incubated with an HRP-conjugated anti-6xHis monoclonal antibody (1:2000, ThermoFisher) for 1 h at 4°C.
After three washing steps with PBS/0.1% (w/v) BSA, bound samples were detected using Ultra-TMB (100 µL/well, ThermoFisher) as the substrate. The reaction was stopped after 15 minutes by addition of 100 µl, 0.2M H₂SO₄, and the absorbance at 450 nm was measured using an Infinite M200PRO plate reader (Tecan, Mannedorf, Switzerland). Data were analyzed with Prism7 (GraphPad Software).

**Cleavage of SNARE proteins in rat brain detergent extracts (BDE):** Rat BDE were prepared from fresh dissected adult rat brains as previously described. Briefly, a rat brain was homogenized in 15 ml 320 mM sucrose buffer, followed by a centrifugation at 5000 rpm for 2 min at 4°C. Supernatants were collected and centrifuged at 11,000 rpm for 12 min. The pellet was collected and solubilized for 30 min in 15 ml Tris-buffered saline (TBS: 20 mM Tris, 150 mM NaCl) plus 2% of Triton X-100 and a cocktail of protease inhibitors (Roche, CA). Samples were subsequently centrifuged at 17,000 rpm for 20 min to remove the insoluble materials. The final BDE concentration is ~ 2 mg/ml proteins. BDE (60 µl) were incubated with X-LC (0.5 µM), A-LC (1 µM), or B-LC (1 µM), respectively, for 1 hour at 37 °C, and were then analyzed by immunoblot using the enhanced chemiluminescence (ECL) method (Pierce). As controls, LCs were pre-incubated with 20 mM EDTA for 20 minutes at room temperature (RT) to de-active their activity prior to adding into BDE.

**Cleavage of recombinant VAMP by X-LC:** VAMP2 (1-96) was expressed and purified as a His6-tagged protein and VAMP2 (33-86) was expressed and purified as a GST-tagged protein. These proteins (0.6 mg/ml) were incubated with 0.1 µM X-LC in TBS buffer for 1 hour at 37 °C. Samples were either analyzed by SDS-PAGE gels and Coomassie Blue staining, or subjected to mass spectrometry analysis.

**Cleavage of VAMPs in cell lysates:** Full-length HA-tagged VAMP1, 3, 7, and 8 were transfected into HEK293 cells using PolyJet transfection reagents (SignaGen, MD) following the manufacturer's instruction. Cell lysates were harvested 48 hours later in RIPA buffer (50 mM Tris, 1% NP40, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 400 µl per 10-cm dish) plus a protease inhibitor cocktail (Sigma-Aldrich). Cell lysates (250 µl) were incubated with X-LC (0.5 µM) for 1 hour at 37 °C. Samples were then analyzed by immunoblot.

**Whole Protein Analysis by LC-MS/MS:** Samples were analyzed at Taplin Biological Mass Spectrometry Core Facility at Harvard Medical School. Briefly, whole protein samples were loaded onto a 100 µm internal diameter C18 reverse phase HPLC column packed with 3cm of beads off-line using a pressure cell. The column was re-attached to an Accela 600 Pump (Thermo Fisher Scientific, Waltham, MA). A rapid gradient of increasing acetonitrile...
was used to elute the protein/peptide from the HPLC column. As peptides eluted, they were subjected to electrospray ionization and then entered into an LTQ Orbitrap Velos Pro ion-trap mass spectrometer (Thermo Fisher Scientific, Waltham, MA) to acquire a high resolution FTMS scan at 60000 resolution, a second scan at low resolution in the ion trap, and a final scan to perform data dependent MS/MS. The charge state envelopes were de-convoluted manually to obtain mono-isotopic masses when possible or average masses for the proteins. Peptide and protein identity were determined by matching protein databases with the acquired fragmentation pattern by the software program, Sequest (Thermo Fisher Scientific). All databases include a reversed version of all the sequences and the data was filtered to between a one and two percent peptide false discovery rate.

**Identification of the protease cleavage site between LC and H₅:** His6-tagged recombinant X-LC-HN fragment (residues 1-891) was purified in E.coli and subjected to limited proteolysis by endoproteinase Lys-C (Sigma P2289, 100:1 (toxin:Lys-C) molar ratio, 25 minutes at room temperature). The cleavage site was determined by Tandem Mass Tag (TMT) labeling and tandem mass spectrometry approach. Briefly, intact X-LC-HN samples were labeled with the light TMT and equal amount of X-LC-H₅ samples were exposed to Lys-C and then labeled with the heavy TMT. Both samples were then digested with chymotrypsin, combined together, and subjected to quantitative mass spectrometry analysis.

**Cysteine alkylation by NEM:** Lys-C activated X-LC-HN fragment was diluted into sodium phosphate buffer (10 mM, pH 6.5) at the final concentration of 0.3 mg/ml, with or without NEM at indicated concentrations (20, 10, and 5 mM) and incubated for 10 minutes at RT. NEM was freshly prepared in sodium phosphate buffer. Samples were mixed with 3x neutral loading dyes (200 mM Tris pH 6.8, 30% glycerol, 6% Lithium Dodecyl sulfate, 10mM NEM and 0.06% BPB). For samples without NEM, the same 3x SDS loading dye without NEM was used. Samples were further incubated with the loading dye at RT for 10 minutes, heated for 10 min at 55 °C, and then analyzed by SDS-PAGE and Coomassie Blue staining.

**Neuron culture and immunoblot analysis:** Primary rat cortical neurons were prepared from E18-19 embryos using a papain dissociation kit (Worthington Biochemical, NJ), as we described previously. Experiments were carried out on DrV 14-16. Neurons were exposed to BoNT/X fragments or sortase ligation mixture in media for 12 hrs. Cells were then washed and lysed with RIPA buffer (50 mM Tris, 1% NP40, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS) plus a protease inhibitor cocktail (Sigma-Aldrich). Lysates were centrifuged for 10
min at maximum speed using a microcentrifuge at 4°C. Supernatants were subjected to SDS-PAG and immunoblot analysis.

**Dot blot:** BoNTs (0.2 μg in 1 μl) were spotted onto nitrocellulose membranes and dried (10 minutes at room temperature). The membranes were blocked with 5% milk in TBST (TBS plus 0.05% Tween20) for 30 min and then incubated with indicated antisera (1:500 dilution) for 30 min. The membranes were then washed three times with TBST and incubated with HRP (horseradish peroxidase) conjugated secondary antibodies for 30 min, washed three more times with TBST, and analyzed with the ECL method (Pierce). We note that the BoNT/X sample was composed of purified X-LC-H\(_N\) and X-34 at 1:1 ratio.

**Sortase-mediated ligation:** GST-X-Hc was cleaved overnight at 4 °C by thrombin before adding into the mixture of proteins. Ligation reaction was set up in 50 μl TBS buffer with addition of X-LC-H\(_N\) (8 μM), thrombin-cleaved GST-X-H\(_C\) (25 μM), Ca\(^{2+}\) (10 mM), and sortase (10 μM), for 40 min at RT. In FIG. 4C, neurons were exposed to 5 μl of the mixture in media for 12 hrs. In DAS assay described in FIG. 4D, 25 μl of the mixture was injected into the hind leg of mice.

**DAS assay:** Sortase ligation mixture was first activated with limited proteolysis using trypsin (60:1 molar ratio (the total amount of the proteins : trypsin), 30 min at RT). We chose trypsin instead of Lys-C here as we can stop the proteolysis by adding trypsin inhibitor (Soybean trypsin inhibitor, 1:10 ratio (trypsin : trypsin inhibitor). Mice (CD-I strain, 21-25g, n = 6) were anesthetized with isoflurane (3-4%) and were injected with sortase ligation mixture using a 30-gauge needle attached to the sterile Hamilton Syringes, into the gastrocnemius muscles of the right hind limb. Botulinum result in paralysis of the hind paw in the startle response. Muscle paralysis was observed within 12 hours after the injection as previously described.

**References**


Binz, T., Bade, S., Rummel, A., Kollewe, A. & Alves, J. Arg(362) and Tyr(365) of the botulinum neurotoxin type a light chain are involved in transition state stabilization. Biochemistry 41, 1717-1723 (2002).


OTHER EMBODIMENTS

All of the features disclosed in this specification may be combined in any combination.
Each feature disclosed in this specification may be replaced by an alternative feature serving the same, equivalent, or similar purpose. Thus, unless expressly stated otherwise, each feature disclosed is only an example of a generic series of equivalent or similar features.

From the above description, one skilled in the art can easily ascertain the essential characteristics of the present disclosure, and without departing from the spirit and scope thereof, can make various changes and modifications of the disclosure to adapt it to various usages and conditions. Thus, other embodiments are also within the claims.

EQUIVALENTS AND SCOPE

While several inventive embodiments have been described and illustrated herein, those of ordinary skill in the art will readily envision a variety of other means and/or structures for performing the function and/or obtaining the results and/or one or more of the advantages described herein, and each of such variations and/or modifications is deemed to be within the scope of the inventive embodiments described herein. More generally, those skilled in the art will readily appreciate that all parameters, dimensions, materials, and configurations described herein are meant to be exemplary and that the actual parameters, dimensions, materials, and/or configurations will depend upon the specific application or applications for which the inventive teachings is/are used. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific inventive embodiments described herein. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that, within the scope of the appended claims and equivalents thereto, inventive embodiments may be practiced otherwise than as specifically described and claimed. Inventive embodiments of the present disclosure are directed to each individual feature, system, article, material, kit, and/or method described herein. In addition, any combination of two or more such features, systems, articles, materials, kits, and/or methods, if such features, systems, articles, materials, kits, and/or methods are not mutually inconsistent, is included within the inventive scope of the present disclosure.

All definitions, as defined and used herein, should be understood to control over dictionary definitions, definitions in documents incorporated by reference, and/or ordinary meanings of the defined terms.

All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.
The indefinite articles "a" and "an," as used herein in the specification and in the 
claims, unless clearly indicated to the contrary, should be understood to mean "at least one."

The phrase "and/or," as used herein in the specification and in the claims, should be 
understood to mean "either or both" of the elements so conjoined, i.e., elements that are 
conjunctively present in some cases and disjunctively present in other cases. Multiple 
elements listed with "and/or" should be construed in the same fashion, i.e., "one or more" of 
the elements so conjoined. Other elements may optionally be present other than the elements 
specifically identified by the "and/or" clause, whether related or unrelated to those elements 
specifically identified. Thus, as a non-limiting example, a reference to "A and/or B", when 
used in conjunction with open-ended language such as "comprising" can refer, in one 
embodiment, to A only (optionally including elements other than B); in another embodiment, 
to B only (optionally including elements other than A); in yet another embodiment, to both A 
and B (optionally including other elements); etc.

As used herein in the specification and in the claims, "or" should be understood to have 
the same meaning as "and/or" as defined above. For example, when separating items in a list, 
"or" or "and/or" shall be interpreted as being inclusive, i.e., the inclusion of at least one, but 
also including more than one, of a number or list of elements, and, optionally, additional 
unlisted items. Only terms clearly indicated to the contrary, such as "only one of" or "exactly 
one of," or, when used in the claims, "consisting of," will refer to the inclusion of exactly one 
element of a number or list of elements. In general, the term "or" as used herein shall only be 
interpreted as indicating exclusive alternatives (i.e. "one or the other but not both") when 
preceded by terms of exclusivity, such as "either," "one of," "only one of," or "exactly one of." 
"Consisting essentially of," when used in the claims, shall have its ordinary meaning as used in 
the field of patent law.

As used herein in the specification and in the claims, the phrase "at least one," in 
reference to a list of one or more elements, should be understood to mean at least one element 
selected from any one or more of the elements in the list of elements, but not necessarily 
including at least one of each and every element specifically listed within the list of elements 
and not excluding any combinations of elements in the list of elements. This definition also 
allows that elements may optionally be present other than the elements specifically identified 
within the list of elements to which the phrase "at least one" refers, whether related or 
unrelated to those elements specifically identified. Thus, as a non-limiting example, "at least 
one of A and B" (or, equivalently, "at least one of A or B," or, equivalently "at least one of A
and/or B") can refer, in one embodiment, to at least one, optionally including more than one, A, with no B present (and optionally including elements other than B); in another embodiment, to at least one, optionally including more than one, B, with no A present (and optionally including elements other than A); in yet another embodiment, to at least one, optionally including more than one, A, and at least one, optionally including more than one, B (and optionally including other elements); etc.

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

In the claims, as well as in the specification above, all transitional phrases such as "comprising," "including," "carrying," "having," "containing," "involving," "holding," "composed of," and the like are to be understood to be open-ended, i.e., to mean including but not limited to. Only the transitional phrases "consisting of" and "consisting essentially of shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03.

What is claimed is:
CLAIMS

1. An isolated Clostridial Botulinum neurotoxin (BoNT) polypeptide comprising the amino acid sequence of SEQ ID NO: 1.

2. An isolated BoNT polypeptide, comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 1.

3. The isolated BoNT polypeptide of claim 1, consisting of the amino acid sequence of SEQ ID NO: 1.

4. An isolated Clostridial Botulinum neurotoxin (BoNT) polypeptide comprising the amino acid sequence of SEQ ID NO: 2.

5. An isolated BoNT polypeptide, comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 2.

6. The isolated BoNT polypeptide of claim 4, consisting of the amino acid sequence of SEQ ID NO: 2.

7. An isolated Clostridial Botulinum neurotoxin (BoNT) polypeptide comprising the amino acid sequence of SEQ ID NO: 3.

8. An isolated BoNT polypeptide, comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to SEQ ID NO: 3.
9. The isolated BoNT polypeptide of claim 7, consisting of the amino acid sequence of SEQ ID NO: 3.

10. A modified Clostridial Botulinum neurotoxin (BoNT) polypeptide, comprising one or more substitution mutation(s) in a position corresponding to C461, C467, or C1240 of SEQ ID NO: 1.


12. The modified BoNT polypeptide of claim 10, comprising the amino acid sequence of any one of SEQ ID NO: 4-17.

13. The modified BoNT polypeptide of claim 10, comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any of SEQ ID NOs: 4-17, wherein the polypeptide does not have the amino acid sequence of SEQ ID NO: 1.

14. The modified BoNT polypeptide of claim 10, consisting of the amino acid sequence of any one of SEQ ID NOs: 4-17.

15. A modified Clostridial Botulinum neurotoxin (BoNT) polypeptide, comprising a single substitution mutation in a position corresponding to C461 or C467 of SEQ ID NO: 2.

17. The modified BoNT polypeptide of claim 15 comprising the amino acid sequence of any one of SEQ ID NOs: 18-21.

18. The modified BoNT polypeptide of claim 15, comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any of SEQ ID NOs: 18-21, wherein the polypeptide does not have the amino acid sequence of SEQ ID NO: 2.

19. The modified BoNT polypeptide of claim 15, consisting of the amino acid sequence of any one of SEQ ID NOs: 18-21.

20. A chimeric *Clostridial Botulinum* neurotoxin (BoNT) polypeptide comprising the amino acid sequence of any one of SEQ ID NOs: 22-24.

21. The chimeric BoNT polypeptide of claim 20, comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any one of SEQ ID NOs: 22-24, wherein the polypeptide does not have the amino acid sequence of SEQ ID NO: 1 or SEQ ID NO: 2.

22. The chimeric BoNT polypeptide of claim 20, consisting of the amino acid sequence of any one of SEQ ID NOs: 22-24.

23. The chimeric BoNT polypeptide of claim 20, further comprising a single substitution mutation in a position corresponding to C461 or C467 of SEQ ID NO: 2.

24. The chimeric BoNT polypeptide of claim 23, comprising the amino acid sequence of any one of SEQ ID NOs: 25-30.

25. The chimeric BoNT polypeptide of claim 23, comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least
91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any one of SEQ ID NOs: 25-30.

26. The chimeric BoNT polypeptide of claim 23, consisting of the amino acid sequence of any one of SEQ ID NOs: 25-30.

27. The BoNT polypeptide of any one of claims 1-26, wherein the BoNT polypeptide enters a cell.

28. The BoNT polypeptide of claim 27, wherein the BoNT polypeptide cleaves a SNARE protein in the cell.

29. The BoNT polypeptide of claim 28, wherein the SNARE protein is selected from the group consisting of: SNAP-25, VAMPI, VAMP2, VAMP3, VAMP4, VAMP5, Ykt6, and syntaxin 1.

30. The BoNT polypeptide of claim 29, wherein the SNARE protein is VAMPI.

31. The BoNT polypeptide of claim 26, wherein the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 39.

32. The BoNT polypeptide of claim 29, wherein the SNARE protein is VAMP2.

33. The BoNT polypeptide of claim 32, wherein the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 40.

34. The BoNT polypeptide of claim 29, wherein the SNARE protein is VAMP3.

35. The BoNT polypeptide of claim 34, wherein the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 41.

36. The BoNT polypeptide of claim 29, wherein the SNARE protein is VAMP4.
37. The BoNT polypeptide of claim 36, wherein the BoNT cleaves between amino acid residues corresponding to K87 and S88 of SEQ ID NO: 42.

38. The BoNT polypeptide of claim 29, wherein the SNARE protein is VAMP5.

39. The BoNT polypeptide of claim 30, wherein the BoNT cleaves between amino acid residues corresponding to R40 and S41 of SEQ ID NO: 43.

40. The BoNT polypeptide of claim 29, wherein the SNARE protein is Ykt6.

41. The BoNT polypeptide of claim 40, wherein the BoNT cleaves between amino acid residues corresponding to K173 and S174 of SEQ ID NO: 44.

42. The BoNT polypeptide of any one of claims 10-19, wherein the BoNT polypeptide has increased stability compared to its corresponding wild type BoNT polypeptide.

43. The BoNT polypeptide of any one of claims 27-42, wherein the cell is a secretory cell.

44. The BoNT polypeptide of claim 43, wherein the cell is a neuronal cell.

45. The BoNT polypeptide of claim 43, wherein the cell is an immune cell.

46. The BoNT polypeptide of claim 45, wherein the BoNT polypeptide suppresses neuronal activity.

47. The BoNT polypeptide of claim 46, wherein the BoNT polypeptide induces flaccid paralysis.

48. The BoNT polypeptide of any one of claims 27 to 47, wherein the cell is a cultured cell.

49. The BoNT polypeptide of any one of claims 27 to 47, wherein the cell is in vivo.
50. The BoNT polypeptide of any one of claims 27 to 49, wherein the cell is from a mammal.

51. The BoNT polypeptide of claim 50, wherein the mammal is a human.

52. The BoNT polypeptide of claim 50, wherein the mammal is a rodent.

53. The BoNT polypeptide of claim 52, wherein the rodent is a mouse.

54. The BoNT polypeptide of claim 52, wherein the rodent is a rat.

55. The BoNT polypeptide of any one of claims 1-54, wherein the BoNT polypeptide does not cross react with an antibody against BoNT serotype A, B, C, D, E, F, or G.

56. A nucleic acid molecule comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%, or 100% identity to the BoNT polypeptide of any one of claims 1-55.

57. A nucleic acid vector comprising the nucleic acid molecule of claim 56.

58. A cell comprising the nucleic acid molecule of claim 56 or the nucleic acid vector of claim 57.

59. A cell expressing the BoNT polypeptide of any one of claims 1-55.

60. A method of producing BoNT polypeptide of any one of claims 1-55, the method comprising the steps of culturing the cell of claim 59 under conditions wherein said BoNT polypeptide is produced.

61. The method of claim 60, further comprising recovering the BoNT polypeptide from the culture.
62. A modified *Clostridial Botulinum* neurotoxin (BoNT) polypeptide comprising:
(a) a protease domain;
(b) a modified linker region; and
(c) a translocation domain;
wherein (a), (b), and (c) are from BoNT serotype X, and wherein the modified linker region comprises one single substitution mutation in a position corresponding to C461 or C467 of SEQ ID NO: 1.

63. The modified BoNT polypeptide of claim 62, further comprising: (d) a receptor binding domain.

64. The modified BoNT polypeptide of claim 62 or 63, wherein the modified linker region comprises a substitution mutation corresponding to C461S or C461A in SEQ ID NO: 1.

65. The modified BoNT polypeptide of claim 62 or 63, wherein the modified linker region comprises a substitution mutation corresponding to C467S or C467A in SEQ ID NO: 1.

66. The modified BoNT polypeptide of claim 63, wherein the receptor binding domain is from BoNT/X.

67. The modified BoNT polypeptide of claim 66, wherein the receptor binding domain is modified.

68. The modified BoNT polypeptide of claim 67, wherein the receptor binding domain comprises a substitution mutation corresponding to C1240S or C1240A in SEQ ID NO: 1.

69. The modified BoNT polypeptide of any one of claims 62-68, wherein the receptor binding domain is from serotype selected from the group consisting of A, B, C, D, E, F, and G.

70. The modified BoNT polypeptide of any one of claims 62-69, wherein the modified BoNT polypeptide enters a cell.
71. The modified BoNT polypeptide of claim 70, wherein the modified BoNT polypeptide cleaves SNARE proteins in the cell.

72. The modified BoNT polypeptide of claim 71, wherein the SNARE protein is selected from the group consisting of: SNAP-25, VAMP1, VAMP2, VAMP3, VAMP4, VAMP5, Ykt6, and syntaxin 1.

73. The modified BoNT polypeptide of claim 72, wherein the SNARE protein is VAMP1.

74. The modified BoNT polypeptide of claim 73, wherein the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 39.

75. The modified BoNT polypeptide of 72, wherein the SNARE protein is VAMP2.

76. The modified BoNT polypeptide of claim 73, wherein the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 40.

77. The modified BoNT polypeptide of claim 72, wherein the SNARE protein is VAMP3.

78. The modified BoNT polypeptide of claim 77, wherein the BoNT cleaves between amino acid residues corresponding to R66 and A67 of SEQ ID NO: 41.

79. The modified BoNT polypeptide of claim 72, wherein the SNARE protein is VAMP4.

80. The BoNT polypeptide of claim 79, wherein the BoNT cleaves between amino acid residues corresponding to K87 and S88 of SEQ ID NO: 42.

81. The modified BoNT polypeptide of claim 72, wherein the SNARE protein is VAMP5.

82. The BoNT polypeptide of claim 81, wherein the BoNT cleaves between amino acid residues corresponding to R40 and S41 of SEQ ID NO: 43.

83. The modified BoNT polypeptide of claim 72, wherein the SNARE protein is Ykt6.
84. The BoNT polypeptide of claim 83, wherein the BoNT cleaves between amino acid residues corresponding to K173 and S174 of SEQ ID NO: 44.

85. The modified BoNT polypeptide of any one of claims 62-84, wherein the modified BoNT polypeptide has increased stability compared to its corresponding wild type BoNT polypeptide.

86. The modified BoNT polypeptide of any one of claim 62-85, wherein the cell is a secretory cell.

87. The modified BoNT polypeptide of claim 86, wherein the cell is a neuronal cell.

88. The modified BoNT polypeptide of claim 86, wherein the cell is an immune cell.

89. The modified BoNT polypeptide of claim 88, wherein the modified BoNT polypeptide suppresses neuronal activity.

90. The modified BoNT polypeptide of claim 89, wherein the modified BoNT polypeptide induces flaccid paralysis.

91. The modified BoNT polypeptide of any one of claims 70-90, wherein the cell is a cultured cell.

92. The modified BoNT polypeptide of any one of claims 70-90, wherein the cell in vivo.

93. The modified BoNT polypeptide of any one of claim 91 or claim 92, wherein the cell is from a mammal.

94. The BoNT polypeptide of claim 93, wherein the mammal is a human.

95. The BoNT polypeptide of claim 93, wherein the mammal is a rodent.
96. The BoNT polypeptide of claim 95, wherein the rodent is a mice.

97. The BoNT polypeptide of claim 95, wherein the rodent is a rat.

98. The modified BoNT polypeptide of any one of claims 62-97, wherein the modified linker region comprises an artificial linker.

99. The modified BoNT polypeptide of claim 98, wherein the artificial linker comprises a cleavage site of a protease.

100. The modified BoNT polypeptide of claim 99, wherein the protease is selected from the group consisting of thrombin, TEV, PreScission (3C protease), Factor Xa, Enterokinase, and SUMO protease.

101. The modified BoNT polypeptide of claim 100, wherein the linker comprises an amino acid sequence of any of SEQ ID NOs: 50-56.

102. A nucleic acid molecule comprising a polynucleotide encoding a modified BoNT polypeptide comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%, or 100% identity to the modified BoNT polypeptide of any one of claims 62-101.

103. A nucleic acid vector comprising the nucleic acid molecule of claim 102.

104. A cell comprising the nucleic acid molecule of claim 102 or the nucleic acid vector of claim B42.

105. A cell expressing the modified BoNT polypeptide of any one of claims 62-101.

106. A method of producing a modified botulinum neurotoxin (BoNT) polypeptide, the method comprising the steps of culturing the cell of claim 105 under conditions wherein said modified BoNT polypeptide is produced.
107. The method of claim 106, further comprising recovering the modified BoNT polypeptide from the culture.

108. A modified *Clostridial Botulinum* neurotoxin (BoNT) polypeptide, comprising one or more substitution mutation(s) in a position corresponding to R360, Y363, H227, E228, or H231 in SEQ ID NO: 1.

109. The modified BoNT polypeptide of claim 108, wherein the one or more substitution mutation(s) corresponds to R360A/Y363F, H227Y, E228Q, or H231Y in SEQ ID NO: 1.

110. The modified BoNT polypeptide of claim 109, comprising the amino acid sequence of any one of SEQ ID NOs: 31-38.

111. The modified BoNT polypeptide of claim 108, comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identity to any of SEQ ID NOs: 31-38, wherein the polypeptide does not have the amino acid sequence of SEQ ID NO: 1 or SEQ ID NO: 2.

112. The modified BoNT polypeptide of claim 110, consisting of the amino acid sequence of any one of SEQ ID NOs: 31-38.

113. A modified *Clostridial Botulinum* neurotoxin, serotype X (BoNT/X) polypeptide comprising:

(a) an inactive protease domain;
(b) a linker region; and
(c) a translocation domain

114. The modified BoNT/X polypeptide of claim 113, wherein the modified BoNT/X further comprises a receptor binding domain.
115. The modified BoNT/X polypeptide of claims 113 or 114, wherein the inactive protease domain comprises one or more substitution mutation(s) in a position corresponding to R360, Y363, H227, E228, or H231 of SEQ ID NO: 1.

116. The modified BoNT polypeptide of claim 115, wherein the one or more substitution mutation(s) corresponds to R360A/Y363F, H227Y, E228Q, or H231Y of SEQ ID NO: 1.

117. The modified BoNT polypeptide of any one of claims 108-116, wherein the modified BoNT polypeptide enters a cell.

118. The modified BoNT/X polypeptide of claim 117, wherein the modified BoNT polypeptide does not cleave a SNARE protein.

119. The modified BoNT/X polypeptide of any one of claims 108-118, further comprising a modification in the linker region of (b).

120. The modified BoNT/X polypeptide of claim 119, wherein the modification in the linker region comprises one single substitution mutation in a position corresponding to C461 or C467 of SEQ ID NO: 1.

121. The modified BoNT/X polypeptide of claim 120, wherein the single substitution mutation corresponds to C461A, C461S, C467A, or C467S in SEQ ID NO: 1.

122. The modified BoNT/X polypeptide of any one of claims 114-121, further comprising a modification in the receptor binding domain of (d).

123. The modified BoNT/X of claim 122, wherein the modification in the receptor binding domain comprises a substitution mutation in a position corresponding to C1240 of SEQ ID NO: 1.

124. A nucleic acid molecule comprising a polynucleotide encoding a modified BoNT polypeptide comprising an amino acid sequence that has at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least
94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%, or 100% identity to the modified polypeptide of any one of claims 108-123.

125. A nucleic acid vector comprising the nucleic acid molecule of claim 124.

126. A cell comprising the nucleic acid molecule of claim C17 or the nucleic acid vector of claim 125.

127. A cell expressing the modified BoNT polypeptide of any one of claims 108-123.

128. A method of producing a modified botulinum neurotoxin (BoNT) polypeptide, the method comprising the steps of culturing the cell of claim 127 under conditions wherein said modified BoNT polypeptide is produced.

129. The method of claim 128, further comprising recovering the modified BoNT polypeptide from the culture.

130. The modified BoNT polypeptide of any one of claims 108-123, for use as a delivery vehicle to deliver therapeutics into neurons.

131. A chimeric molecule comprising a first portion linked to a second portion, wherein the first portion is a modified BoNT polypeptide of any one of claims 108-123.

132. The chimeric molecule of any one of claims 131, wherein the first portion and the second portion are linked covalently.

133. The chimeric molecule of claim 131, wherein the first portion and the second portion are linked non-covalently.

134. The chimeric molecule of any one of claims 131-133, wherein the second portion is selected from the group consisting of a small molecule, a nucleic acid, a short polypeptide and a protein.
135. The chimeric molecule of claim 134, wherein the second portion is a bioactive molecule.

136. The chimeric molecule of claim 134 or claim 135, wherein the second portion is a non-poly-peptide drug.

137. The chimeric molecule of claims 134 or claim 135, wherein the second portion is a therapeutic polypeptide.

138. A nucleic acid molecule comprising a polynucleotide sequence that encodes the chimeric molecule of claim 137.

139. A nucleic acid vector comprising the nucleic acid molecule of claim 138.

140. A cell comprising the nucleic acid molecule of claim 138 or the nucleic acid vector of claim 139.

141. A cell expressing the chimeric molecule of claim 140.

142. A method of producing a chimeric molecule, the method comprising the steps of culturing the cell of claim 141 under conditions wherein said chimeric molecule is produced.

143. The method of claim 142, further comprising recovering the chimeric molecule from the culture.

144. A pharmaceutical composition comprising the BoNT polypeptide of any one of claims 1-55, 62-101, or 108-123.

145. A pharmaceutical composition comprising the chimeric molecule of any one of claims 131-137.

146. The pharmaceutical composition of claims 144 or 145, further comprising a pharmaceutically acceptable excipient.
147. A kit comprising a pharmaceutical composition of any one of claims 144-146 and directions for therapeutic administration of the pharmaceutical composition.

148. A method of treating a condition, the method comprising administering a therapeutically effective amount of the modified BoNT polypeptide of any one of claims 1-55, 62-101, or 108-123, or the chimeric molecule of any one of claims 131-137, or the pharmaceutical composition of any one of claims 144-146 to a subject to treat the condition.

149. The method of claim 148, wherein the condition is associated with overactive neurons or glands.

150. The method of claim 148 or claim 149, wherein the condition is selected from the group consisting of: spasmodic dysphonia, spasmodic torticollis, laryngeal dystonia, oromandibular dysphonia, lingual dystonia, cervical dystonia, focal hand dystonia, blepharospasm, strabismus, hemifacial spasm, eyelid disorder, cerebral palsy, focal spasticity and other voice disorders, spasmodic colitis, neurogenic bladder, anismus, limb spasticity, tics, tremors, bruxism, anal fissure, achalasia, dysphagia and other muscle tone disorders and other disorders characterized by involuntary movements of muscle groups, lacrimation, hyperhydrosis, excessive salivation, excessive gastrointestinal secretions, secretory disorders, pain from muscle spasms, headache pain, dermatological or aesthetic/cosmetic conditions, obesity/reduced appetite.

151. The method of claim 150, wherein the condition is not associated with unwanted neuronal activity.

152. The method of claim 151, wherein the condition is selected from the group consisting of: psoriasis, allergy, haemophagocytic lymphohistiocytosis, and alcoholic pancreatic disease.

153. The method of any one of claims claim F1-F5, wherein the administering is via injection to where unwanted neuronal activity is present.
154. The modified BoNT polypeptide of any one of claims 1-55, 62-101, 108-123, or the chimeric molecule of any one of claims 131-137, or the pharmaceutical composition of claims 144-146, for use in treating a condition associated with unwanted neuronal activity.

155. The modified BoNT polypeptide of any one of claims 1-55, 62-101, 108-123, or the chimeric molecule of any one of claims 131-137, or the pharmaceutical composition of claims 144-146, for use in medicine.

156. A method of producing a Clostridial Botulinum neurotoxin (BoNT) polypeptide, the method comprising:
   (i) obtaining a first BoNT fragment comprising a light chain (LC) and a N-terminal domain of a heavy chain (¾), wherein the first BoNT fragment comprises a C-terminal LPXTGG (SEQ ID NO: 60) motif;
   (ii) obtaining a second BoNT fragment comprising a C-terminal domain of the heavy chain (He); wherein the second BoNT fragment comprise a specific protease cleavage site at its N-terminus;
   (iii) cleaving the second BoNT fragment with a specific protease, wherein the cleavage results in a free Glycine residue at the N-terminus; and
   (iv) contacting the first BoNT fragment and the second BoNT fragment in the presence of a transpeptidase, thereby ligating the first BoNT fragment and the second BoNT fragment to form a ligated BoNT.

157. The method of claim 156, wherein the first BoNT fragment further comprises an affinity tag.

158. The method of claim 157, wherein the affinity tag is fused to the first BoNT fragment at the N-terminus.

159. The method of claim 157, wherein the affinity tag is fused to the first BoNT fragment at the C-terminus.
160. The method of any one of claims 156-159, wherein the affinity tag is selected from the group consisting of: His6, GST, Avi, Strep, S, MBP, Sumo, FLAG, HA, Myc, SBP, E, Calmodulin, Softag 1, Softag 3, TC, V5, VSV, Xpress, Halo, and Fc.

161. The method of any one of claims 156-160, wherein the second BoNT fragment further comprises an affinity tag.

162. The method of claim 161, wherein the affinity tag is fused to the second BoNT fragment at the N-terminus.

163. The method of claim 161, wherein the affinity tag is fused to the second BoNT fragment at the C-terminus.

164. The method of any one of claims 161-163, wherein the affinity tag is selected from the group consisting of: His6, GST, Avi, Strep, S, MBP, Sumo, FLAG, HA, Myc, SBP, E, Calmodulin, Softag 1, Softag 3, TC, V5, VSV, Xpress, Halo, and Fc.

165. The method of any one of claims 161-164, wherein the protease is selected from the group consisting of: thrombin, TEV, PreScission (3C protease), Enterokinase, and SUMO protease.

166. The method of claim 165, wherein the cognate protease is thrombin.

167. The method of any one of claims 156-166, wherein the first BoNT fragment is from BoNT serotype A, B, C, D, E, F, G, or X.

168. The method of claim 167, wherein the first BoNT fragment is from BoNT/X.

169. The method of any one of claims 156-168, wherein the second BoNT fragment is from BoNT serotype A, B, C, D, E, F, G, or X.

170. The method of claim 169, wherein the second BoNT fragment is from BoNT/A.
171. The method of claim 169, wherein the second BoNT fragment is from BoNT/B.

172. The method claim 169, wherein the second BoNT fragment is from BoNT/C.

173. The method of claim 169, wherein the second BoNT fragment is from BoNT/X.

174. The method of any one of claims 156-173, wherein the transpeptidase is a sortase.

175. The method of claim 174, wherein the sortase is from Staphylococcus aureus (SrtA).
FIG. 6

**HPLC**

GGSHHHHHHGMASMTGGQQMGRLYDD
DDKDRWGSMSATAATVPPAAPAGEQGPPA
PPPNTSNRLQQTQAQVDEVVDIMRVNV
DKLIERDKLSDLDDDADALQAGASQFETSA
AKLKRKYWWKNL
FIG. 8A

<table>
<thead>
<tr>
<th></th>
<th>-DTT</th>
<th>+DTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-LC-Hₙ</td>
<td>- + - + + +</td>
<td>- + - + + +</td>
</tr>
<tr>
<td>GST-A-Hₖ</td>
<td>- - - + + +</td>
<td>- - + + + +</td>
</tr>
<tr>
<td>Thrombin</td>
<td>- - + + + +</td>
<td>- - + + + +</td>
</tr>
<tr>
<td>Sortase</td>
<td>- - + + - +</td>
<td>- - + + - +</td>
</tr>
</tbody>
</table>

130kDa = XA-FL
100kDa = X-LC-Hₙ
70kDa = A-Hₖ
55kDa =
40kDa =

FIG. 8B

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X-LC-Hₙ</td>
<td>+ - + + +</td>
</tr>
<tr>
<td>GST-A-Hₖ</td>
<td>- + + + +</td>
</tr>
<tr>
<td>Thrombin</td>
<td>+ + + + +</td>
</tr>
<tr>
<td>Sortase</td>
<td>+ + + - +</td>
</tr>
</tbody>
</table>

actin
syntaxin 1
SNAP-25
VAMP2
### FIG. 9

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>C461S</th>
<th>C467S</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>X-LC-N</strong></td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><strong>X-H₃(C1240S)</strong></td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><strong>sortase</strong></td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><strong>actin</strong></td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><strong>syntaxin 1</strong></td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><strong>SNAP25</strong></td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><strong>VAMP2</strong></td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
FIG. 11A

<table>
<thead>
<tr>
<th></th>
<th>-</th>
<th>2</th>
<th>20</th>
<th>+</th>
<th>2</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lys-C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BoNT/X&lt;sub&gt;RY&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>syntaxin 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNAP-25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAMP2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIG. 11B

FIG. 11C